

What happens after the crash?

The US budget is almost certain to be cut by now. What happens next rests with the reserve banks. Their best plan is to extend credit to the United States, the world's largest debtor.

WASHINGTON'S will to put the United States on a sounder economic footing seems to be wilting as memories of the recent crash on Wall Street fade. The US government has only until tomorrow, 20 November, to avoid the automatic imposition of spending cuts amounting to \$23,000 million under the provisions of the modified Gramm-Rudman Act, and seems bent on prolonging until the last minute its four-week internal wrangle about a more intelligent way of reducing the federal deficit.

Part of the trouble is that the next presidential election campaign is already well under way, with the result that neither party wishes to be seen responsible for either higher taxes or reduced spending (except that the Democrats would cheerfully cut the defence budget). It also seems to be the case that the Congress and the administration have lost heart by recognizing that the object of the exercise is in part symbolic: the danger that there will be an economic recession argues that this is not the time to cut the federal deficit (see *Nature* 330, 1; 1987), but the general opinion has emerged that foreigners will not continue to finance the US trade deficit (witness the falling US dollar) without a demonstration that the United States can quickly come to a simple decision on the federal deficit. It seems not to be fully appreciated elsewhere that the tortuous negotiations between the White House and the Congress are aimed at an outline agreement only, and that it will fall to the appropriations committees to decide the details (and to backslide, given half a chance). That is why Gramm-Rudman may be the best solution.

Meanwhile, the next step in the resolution of the incipient financial crisis needs more attention than it has been given. The major trading partners of the United States, who call themselves the Group of Seven (or G7 for short) have agreed that there should be an international discussion on the aftermath of the stock-market crash, but only when the deficit negotiations in Washington are finished — probably next week. There must be many among the central bankers and finance ministers present who will wish that one among them were nimble enough to pull a financial rabbit out of a hat in the manner of J. M. Keynes, whose proposals at the Bretton Woods conference held sway for more than a quarter of a century (until 1972). Alas, Keynes has been overtaken by events, especially the financial interdependence of industrialized states.

The problem for the next G7 meeting is simply stated. The industrial economies operate factories, farms and service industries for whose products there are potential customers in plenty. For most of the past decade, potential customers have also had the funds with which to make purchases on a scale sufficient to enlarge economic activity everywhere, if unevenly. The obvious snag is that some customers have been purchasing with money borrowed from elsewhere; US customers of the international economy, in particular, have been spending the savings of prudent Japanese and, to a lesser extent, prudent West Germans. Some of the lenders have ended up owning physical assets in the United States, and may not be too worried now, but others, with the US dollar now sharply devalued, know they can get their money back only in kind; if they want cash, their loans will not be fully required. So they have turned shy.

The need, in these circumstances, is to find a mechanism for

sustaining the global demand for goods and services while accommodating the peculiar circumstances of the United States (which must nevertheless eventually balance its books either by increased taxes or by increased voluntary savings). The solution, if there is one, must be a novel device by which governments internationally make up for the unwillingness of private individuals and institutions to finance the trade imbalances that have accumulated, and which cannot suddenly be halted except traumatically. The central banks must find a way to lend each other funds.

In principle, this is what Keynes's International Monetary Fund is for, but the ground rules are too restrictive for present circumstances, while the membership is inappropriately wide. A better instrument is the Bank for International Settlements (BIS), G7's own clearing bank, which exists in Basle, Switzerland, largely as a paper institution for settling reserve banks' debts with each other. Why not turn BIS into an institution that functions fully as an international reserve bank, the lender of last resort? The debtor nations in particular would then find their freedom to execute monetary policy compromised by the terms on which they were compelled to borrow. Either interest rates or currency conversions would be less easily controlled than at present. But that would be better than the prospect of another long recession which everybody seems anxious to avoid. Certainly it would be a more substantial contribution to stability than the agreement, earlier this year, to sustain the US dollar at a level too high for logic to accommodate. □

British research muddle

Are British arrangements for science policy constitutionally improper?

WHAT is to be made of a government that allows one group of taxpayers to purchase the right to influence the spending of other people's taxes? Not much, it will be thought. But that, on the face of things, is what is about to happen with the British government's encouragement of the new Centre for the Exploitation of Science and Technology (see p. 196), to which eighteen industrial companies have made financial contributions. For although the organizers of the new centre insist that they do not wish to force Britain's beleaguered academic researchers into more evidently applied research, there is every reason to fear that the new centre's work will shape the overall pattern of research supported by the research councils, the government's chief source of discretionary expenditure on basic research in academic institutions. The constitutional propriety of this arrangement needs more attention than it has been given. There would be ructions if a committee appointed by, say, insurance companies were to be given an inside track in determining the pattern of spending on the British National Health Service.

This is how malign influence could be exerted. The centre's objective is to survey the developing field of science and to draw attention to fields appearing to offer industrial opportunities. Much will depend on the character and calibre of the people who will eventually staff the centre, but the occasion could well arise



when the group proclaims that the British toothpaste industry, say, has much to gain from a better understanding of non-newtonian liquids (which may well be true, given the meagre literature on the rheology of toothpaste). Taking advantage of its "close links" with the new Advisory Council for Science and Technology (ACOST), the centre will no doubt easily win a ringing declaration that national energies should be bent to the refinement of the extraction of toothpaste from the tubes containing it, whereupon the research councils will listen carefully, marking up research proposals in all cognate fields.

There will no doubt also be more serious issues to be dealt with, such as the utility of high-energy physics as a stimulant of industry. The new centre will not be a going concern until after the British government has decided whether or not to pull out of CERN (the European high-energy physics consortium), but the scope of its opinion on the subject is easily guessed at. (It would point out that high-energy physics is not a profitable field for industrial investment, but might surprise some by noting that British manufacturers should be more energetic in the competition for CERN contracts and might even say that learning more about the processing of particle beams could serve British industry well, not merely in the competition for subcontracts for the US Strategic Defense Initiative.) Whatever the centre's opinion, that is likely to be amplified by ACOST, of which the prime minister is occasionally to be the chairman and on which academic research is represented only in the most formal way. And while the research councils are titularly autonomous, the penalty of ignoring what ACOST says is easily imagined; failure to bend before the wind will risk untried ACOST's disapproval and the possible reduction of the whole science budget that could follow. If that is how it turns out, 18 companies will each have won the right to shape the spending of £600 million for a mere £50,000, a leverage (as the phrase goes) of 10,000 to one.

The plain truth is that the new centre, as advertised, will not be doing the job that British industry needs most urgently to be done. Motor manufacturers continue to spend less on innovative research than their overseas competitors. Electronics manufacturers look primarily to markets opened up by electronics manufacturers elsewhere. Even biotechnology, blessed enthusiastically by the British government ten years ago, seems to play only a marginal part in industrial development. What British industry now needs is what it has needed for most of the past 50 years — constant reminders of the well-known opportunities it has consistently neglected, and persuasive demonstrations of how much profit it has lost thereby. It would be easier to hope that this lesson would occasionally be read by the new administrative apparatus if those concerned were less obviously skewed in their interests towards those of industry. □

Clouds over *perestroika*

What Soviet perestroika most needs now is a demonstration of success. Here's where to start.

MR Mikhail Gorbachev's most urgent need in the pursuit of *perestroika* may be for a stroke of luck. That is the kindest reading of the events of the past few days in Moscow. Mr Gorbachev's difficulty is no different now from what it has been for the past two years, that of persuading all his compatriots, including conservatives and those with a vested interest in the present, that the time has come for change. Yet little seems to be going right for him. His address on the 70th anniversary of the October Revolution, outspoken though it may have been, was a less open repudiation of past domestic abuses of Soviet power than the intellectual community had been hoping for. The sad case of Mr Boris Yeltzin, dismissed for hot-headedness last week from his job as Moscow Party chief, is another sign of the care with which Mr Gorbachev must trim his sails. Now there is the disappointing news that the rate of growth of industrial production in the Soviet Union is lagging behind the target of 4.4

per cent set for this year, which will provide comfort for those eager to say that *perestroika* is not working. The harvest has also been disappointing, although the weather and not *perestroika* is probably to blame. But, Mr Gorbachev may be asking, will nothing go right?

Although, like the rest of us, politicians cannot rely on good luck, they are well-placed to increase the chances when it falls their way. Although Mr Gorbachev's programme for *perestroika* is as much social and administrative as economic, the best place to find the demonstration that he needs that reform can be effective is in the economic field, and in the connection between research and industry. The most obvious difficulty is that much of the research the Soviet government generously supports yields little economic benefit because it runs into the sands of the all-pervading indifference of the bureaucracy. The repeated appeals that Soviet researchers should redouble their efforts to turn discovery to practical application will remain irrelevant so long as there is no particular reason why industrial ministries and their factory managers should disturb themselves by making novel products or by striving after the quality of production familiar in the West. Yet there are several points in the pattern of Soviet research at which it might with advantage be possible to organize the nexus between research and industry so as to prove a kind of existence theorem for *perestroika*. Why not see what might be made of one or two of them?

Instrumentation would be a good starting-point for a demonstration of success. Last month's brief survey of Soviet science (*Nature* 329, 776; 1987) told some tales of groups of talented people working on the development of instruments which are either novel or simpler (and thus potentially cheaper) than those available in the West. There could have been many others — photomultipliers developed at Serpukhov for use in high-energy physics, for example, and high-resolution cathode-ray tubes developed at the Lebedev Physical Institute. The Soviet government appears to have recognized that this is an opportunity to be seized by setting up what it calls an inter-agency complex to look after the development and manufacture of novel scientific instruments, but the mechanism is too cumbersome and its goals at once too grand and insufficiently ambitious. The managers of the complex must still persuade production units to manufacture novel things, which goes against the grain, but are stuck with the hopeless task of satisfying the domestic demand for novel instruments. Meanwhile, they are prompted by their awareness of their ignorance of the worldwide market and their depressed expectations of what Soviet industry can accomplish to talk to manufacturers in the West about the licensing of their more imaginative proposals. Everything goes slowly.

The Soviet interest and the need to demonstrate that *perestroika* can work argue for a different and complementary way of working. If the Soviet economy is eventually to hold its place in the world, it will have to learn to compete in worldwide markets not merely in ingenuity but in the quality and cost of its production. It would make sense that the inter-agency complex for scientific instrumentation should be required to find a way of doing just this for a handful of the instruments (chosen by itself and not by a committee) now in its portfolio, and that it should be given the resources to find out what will sell outside the Soviet Union and to manufacture what it decides to make unhampered by the bureaucracy, most simply by recruiting a new labour force to a newly built manufacturing facility. The general principle is that the Soviet government should find a few places in its economy where the connection between research and production is as direct as in instrumentation, and that it should plead the need to compete with enterprises in the West to rid those concerned from the administrative shackles that cripple its exclusively domestic industry. It should not take long for success to be apparent on a narrow front. The scientific community in the Soviet Union, which has the most to gain from *perestroika*, should bend its intellectual energies to the advocacy of such a course of action. □

Britain's space future in the past as ESA forges on

- "Not a penny more" says UK minister
- ESA decides on long-term programmes

Paris

MINISTERS from West European countries meeting at The Hague on 9 and 10 November cast aside objections from the British government and voted to support a \$12,800 million manned space programme until the end of the century. Mr Kenneth Clarke, Britain's Minister of State for Trade and Industry, had expected to find support from other European Space Agency (ESA) members (see *Nature* 329, 660; 1987) but, when a decision had to be made, Britain was the only country not to agree to the proposals.

"Mr Clarke's position was at least clear", said an ESA spokesman in Paris. Britain will continue to contribute 12.6 per cent (about £85 million) to ESA's overall programme, but, in Clarke's words, "will not give a penny more".

The major elements of the long-term space plan, defined at the last ESA council meeting in Rome in 1985, are the Ariane 5 heavy-lift rocket launcher, the Hermes shuttle and the Columbus 'package', comprising an attached pressurized module for the US space station, a man-tended free-flying laboratory and a polar orbiting platform. Ministers from 12 of the 13 ESA member states gave the go-ahead for Ariane 5 and approved three years of development work on Columbus and Hermes.

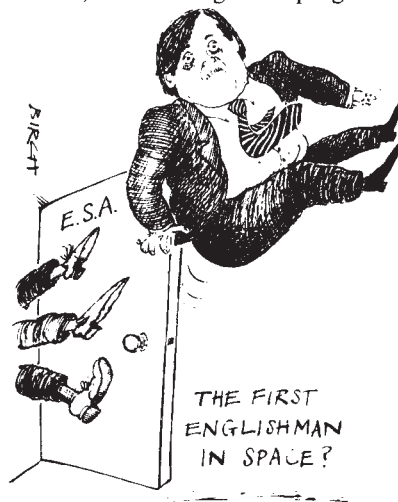
The additional cost of these three projects will mean a 5 per cent annual increase in ESA spending, rising from about \$1,440 million this year to almost \$2,500 million by 1994. Nevertheless, with the exception of Columbus, the projects can be financed almost entirely without Britain's support. Between 80 and 90 per cent of the cost of the three elements has been pledged by France, West Germany and Italy. The French-designed Hermes, criticized by Clarke as "expensive and unnecessary" and "an imitation of a 20-year-old American idea", was oversubscribed by ESA members by 2–6 per cent.

West Germany, ESA's second largest contributor after France, signalled its probable approval of the long-term space plan before the meeting, but tabled a set of provisos. These included a 15–20 per cent cut in the total budget planned for the next 13 years, a three-year postponement of a final decision on Hermes, allowing time for its viability to be proved, and a 'wait and see' approach to Columbus.

As it stands, Columbus is conceived as a part of NASA's space station, but ESA is still not satisfied with the US position with

regard to military usage of the station, European sovereignty in the design, construction and use of APM, ownership of new materials manufactured in its laboratories and arbitration procedures. Members want to await the outcome of negotiations with the United States before committing themselves definitely to the Columbus programme.

Members accepted the provisos and the ESA executive is now looking at ways to make the required savings. One possibility, according to Reimar Lüst, director-general of ESA, would be to delay the timetable for some projects. There is some concern among space scientists, however, that ESA's general programme,



which includes telecommunications satellites, microgravity experiments and Earth observation projects, will be squeezed. Britain is a major contributor to this, mandatory, part of ESA's programme and Clarke says that it will remain the most commercially viable sector. He nevertheless exercised his right of veto to prevent a 5 per cent increase in ESA's basic science programme for 1989, but there are hopes within the agency that he may soften before the budget is finalized.

Work on Ariane 5 and the first phase of Hermes will start right away. The first tests of Ariane's rocket motors are scheduled for 1990 and, by that time, major design obstacles must have been solved for Hermes to be adopted. It is then that members will be asked for the most significant increases in annual contributions.

Now that Britain has opted out of a significant role in ESA's long-term plan, questions are being raised about its continued membership.

Peter Coles

NASA's load to get heavier

Washington

THE National Aeronautics and Space Administration (NASA) announced last week that it was raising the maximum landing weight for future shuttle flights by some 19,000 lb. The change is a result of a three-year analysis of shuttle performance characteristics, and NASA officials are confident that it leaves sufficient safety margins for shuttle operations.

NASA's life sciences division is the most enthusiastic about the higher landing weight, as it will mean that the duration of Spacelab missions can be doubled. Under the old weight allowances, Spacelab was limited to approximately 5 days in orbit before exhausting its fuel cells. Now, sufficient fuel can be carried for a 10-day mission, without exceeding safety margins should a mission need to be aborted.

Landing weight limits are based primarily on analyses of inertial and thermal loads on the shuttle. Gary Coultas, assistant manager of the orbiter project office at the Johnson Space Center in Houston, says that when his office compared design baselines with actual performance data, it became clear that the shuttle could handle heavier loads on landing. For example, the shuttle is designed to withstand forces of 2.5 times gravity during descent, but most shuttle manoeuvres have generated stresses about 1.5–1.6g. As maximum stress loads are lower than expected, heavier payloads can be safely returned to Earth. Coultas says NASA now reckons the shuttle will never need to exceed 2.3g.

The shuttle had been operating with a maximum landing weight of 211,000–214,000 lb. For future missions, that will increase to 230,000 lb. A bigger problem will be aerodynamic stress associated with trajectories needed to launch heavier loads. The first shuttle flight to use the heavier landing weight will be the ASTRO-1 Ultraviolet Astronomical Telescope, now due for launch in 1989.

Shuttle flights are scheduled to resume next summer, although the hoped-for June launch date appears to be slipping. There will be only four missions in 1988; two will launch tracking and data-relay satellites, and two are Defense Department classified missions. The first scheduled scientific payload will be the Magellan Venus radar mapper, now set for an April 1989 launch. Also scheduled for the shuttle in 1989 are the Galileo mission to Jupiter, ASTRO-1, and the Hubble Space Telescope. The first major scientific mission to be launched by NASA in the post-Challenger era will not fly aboard the shuttle at all. The Cosmic Background Explorer will be launched by a Delta rocket in February 1989. Joseph Palca

Research councils struggling in straitjacket of government rules

London

THE system of science funding in Britain is hampering the efforts of the research councils to make more efficient use of scarce resources in the face of declining government support. So much was made clear last week with the publication of the annual reports of the research councils for science and engineering (SERC), the natural environment (NERC) and agriculture and food (AFRC). Although the heads of the councils were keen to point out the year's achievements, the reports contain little to suggest that a revitalization of academic science is in sight.

SERC chairman Professor Bill Mitchell describes 1986-87 as a year of "major achievements and some frustrations". One of the persistent frustrations is the council's vulnerability to sudden large fluctuations in funding demands that are outside its control. Last year's fall in the value of sterling left SERC with a shortfall of £20 million for international subscriptions, while a high pay settlement for scientific staff will cost SERC an extra £10.2 million in the current year. Government provided the extra money, but the system is not conducive to good management practice, says Mitchell.

Another frustration was the council's failure to obtain full international partners for the spallation neutron source ISIS at the Rutherford Appleton Laboratory, which has put in serious doubt Britain's ability to become a member of the European Synchrotron Radiation Facility (see page 198 for more synchrotron news). SERC cannot afford the subscription on current budget forecasts, and the decision now rests with the government.

Mitchell made clear his enthusiasm for interdisciplinary university research centres, but berated the Advisory Board for the Research Councils (ABRC), whose recent advice to the government on the strategic organization of science called for the proportion of SERC's contribution to basic, curiosity-driven science to be 20-30 per cent of its total spend. Mitchell is adamant that a figure of 35-40 per cent, with the remainder on strategic programmes, is the minimum necessary to maintain a viable science base. He accused ABRC of being "apparently prepared to acquiesce in the erosion of basic science". Mitchell's 60-40 split between strategic and basic research precludes an annual subscription to CERN (the European Organization for Nuclear Research) of the present £55 million, from SERC's total budget of £350 million. CERN member states are at present considering recommendations that would reduce administration costs by around 15 per

cent. Even if Britain agrees that such a saving is sufficient to secure its continued membership, other member states may well prefer to see the money reinvested rather than cut from the budget.

For NERC, increased dependence on funds from contract work is causing problems. The bulk of NERC's commissioned research comes from government departments which paid a total of £25 million to the council last year. NERC's complaint is that contracts arise only sporadically and are usually short-term, linked to immediate problem-solving. According to NERC chairman Hugh Fish, research teams need to be maintained during the periods between contracts, with the result that NERC effectively subsidizes the work. Further, under Treasury rules, NERC is prohibited from making a profit on the work. Fish claims that "some departments are not now prepared to pay the full economic costs of commissioned work . . . in effect NERC is forced as a more or less captive contractor to reduce its contract prices to levels which, by all the normal operating rules, are unacceptable".

Income from services to industry approached £10 million, a threefold increase on 1983-84 figures. But NERC has shelved plans to create a company in the private sector to take over its commercial activity: government accounting practices do not allow sufficient management flexibility to run a business effectively.

For AFRC a turbulent period of restructuring is approaching completion, with its former 27 institutions being regrouped into eight. During the year, 401 permanent posts were lost, with 150 compulsory redundancies, bringing the total workforce to 5,275. Earnings from external sources reached £10 million, the target set for 1990. The Ministry of Agriculture says agricultural research will be reduced by £5 million in 1989-90 and £10 million in 1990-91; part of this will almost certainly hit AFRC. AFRC's acting secretary, Professor John Hearn, says the council will be unable to withstand further cuts without damaging its scientific capability beyond repair: "industry will not invest in a lacerated research service".

As AFRC seeks more money from commercial sources, its long-standing arrangement with the Agricultural Genetics Company (AGC) is coming under scrutiny. AGC, a private company, was set up in 1983 and given the first option to exploit certain plant biotechnology discoveries made by AFRC. Difficulties with the relationship arise when outside parties approach AFRC with contracts to work in areas that fall under the AGC agreement.

Simon Hadlington

New centre to exploit research

London

A POTENTIALLY influential new body charged with alerting Britain's government, industry and research community to emerging areas of exploitable technology was formally launched last week. The Centre for Exploitation of Science and Technology (CEST) will be located at the Manchester Science Park, adjacent to the University of Manchester. Initial funding, for a five-year period, has been raised from 18 founding companies, including ICI, Rolls-Royce and Smiths Industries, each contributing £250,000. The taxpayer has given a further £1 million.

The idea for the centre grew from a report by the government's Advisory Committee on Applied Research and Development (ACARD, which has since broadened its remit and been renamed the Advisory Council for Science and Technology, ACOST), which concluded that a forum was needed for research planning (*Nature* 321, 105; 1986).

A chief executive for the centre has yet to be appointed, but Sir Robin Nicholson, chairman of the CEST steering committee and former science adviser in the Cabinet Office, predicts that the centre will be functioning before the end of the year.

CEST will have a core group of about 12 professional staff who will, says Nicholson, spend their time assessing, monitoring and evaluating scientific research worldwide. They will collate information and produce "very high quality reports in specific areas of science and technology and specific areas of industrial need".

Mr John Fairclough, the present science adviser in the Cabinet Office, emphasized the centre's importance in a new "climate of change" where traditional barriers between industrialists and academics should be removed. Both Fairclough and Nicholson insist that the centre's formation is not an attempt to force universities to do more applied research. Nicholson suspects that as the centre's information becomes widely available, industry will recognize areas of strategic importance and respond spontaneously.

CEST will have the ear of the government through its links with ACOST. Academic representation on the steering committee is in the person of Sir David Phillips, chairman of the Advisory Board for the Research Councils.

Thirteen institutions were invited to bid to house the new centre. The successful proposal came from a consortium of nine universities and polytechnics in the North-West. CEST will have access to the consortium's library and computer facilities, as well as considerable academic support.

Simon Hadlington

New fault system to add to Los Angeles earthquake fears

San Francisco

THE earthquake that rocked Los Angeles last month has revealed buried faults running under the city, raising the spectre of potential disaster should more dramatic movements along the fault line occur.

Although the presence of such faults had been hypothesized, the first hard evidence came when data analysis showed that the earthquake on 1 October occurred on a previously unknown fault, rather than the Whittier fault as was first thought. This discovery will require a reevaluation of earthquake risk for the area.

Seismologists and geologists analysing the October earthquake originally placed it at the northernmost end of the Whittier fault (see *Nature* 329, 479; 1987). Like the San Andreas, the Whittier is a strike-slip fault — a crack along which two plates are slipping past each other. But further analysis of the seismic readings showed that the jolt on 1 October came from an east-west thrust fault — a fault along which one plate is slipping under the other. In this case, the block to the north of the fault was forced up and over the southern block, raising the overlying surface by 2 inches.

Kate Hutton, a seismologist at the California Institute of Technology, said that the major aftershock came from a compensatory slippage along a north-south fault lying next to the newly discovered thrust fault. A good candidate for the site of this movement is the northernmost tip of the Whittier fault.

Consulting geologist Thom Davis has been studying the Los Angeles basin for the past year, on contract for the US Geological Survey. Based on his observations of surface hills and folds, he predicted a

system of buried thrust faults extending under Los Angeles and out to the coast. October's earthquake suggests he may be right.

Davis said that the faults are an extension to the south of a known thrust fault system that created the Santa Monica mountains, to the north of Los Angeles. That extension is significant, because it underlies the city of Los Angeles.

To assess the risk of a major earthquake from this fault system, Davis said several types of measurements are needed, including an analysis of the speed of the shortening driving the thrust, and a review of seismic data to see how much of that shortening has been relieved by earthquakes. Some relief may also be derived from 'seismic creep', a steady slippage that builds up no strain.

But if seismic creep and small earthquakes do not add up to enough relief for the measured shortening, the fault system could imply a large earthquake waiting to happen.

Data analysis for the 1 October earthquake has not only changed its origin, but demoted it from a magnitude of 6.1 to 5.9, a 58 per cent reduction in the amount of energy released, said Hutton.

The original 6.1 figure was an estimate based on reports from only a few stations. The aftershock was also downgraded, from 5.5 to 5.3.

Hutton said that the October earthquake, which crumbled a number of old concrete buildings in Whittier, provides little prediction of what kind of damage to expect from a large earthquake, but will be useful in designing buildings better able to resist the medium-sized tremors that occasionally rock the Los Angeles area.

Marcia Barinaga

DNA fingerprint ruling

A MAN has been convicted in a British court this week on the strength of evidence from the 'genetic fingerprint' method pioneered by Dr Alec Jeffreys at Leicester University and published in *Nature* (317, 818; 1985).

Robert Melias was sentenced to 8 years imprisonment for a rape committed in January this year. Bristol Crown Court heard that the match of his DNA profile to that of DNA samples taken from the victim indicates odds of 4,000,000 to one against his being innocent. The conviction has been interpreted by the Home Office as a significant advance in solving crime. In another recent case the test was used to clear a man accused of murder and forensic scientists are now cooperating with police on more than 50 cases involving the fingerprint method. Demand for the fingerprinting kits, produced by ICI subsidiary Cellmark Diagnostics at a price of £120.75, is expected to increase rapidly now that police forces have seen the potential of the method realized.

S.J.H.

Ariane flight delay

FLIGHT 21 of Europe's Ariane 3 commercial rocket, scheduled for 30 December, has been put back by 5 or 6 weeks. Ariane-space, the company responsible for Ariane launches, refused to accept delivery of the rocket's third-stage liquid hydrogen-oxygen motor from Société Européenne de Propulsion (SEP), following concern at performance of the turbopump. Overheating of the pump caused the abortion of flight 18 in May 1986. Flight 20, scheduled for 18 November, will still go ahead. P.C.

AIDS virus cultured

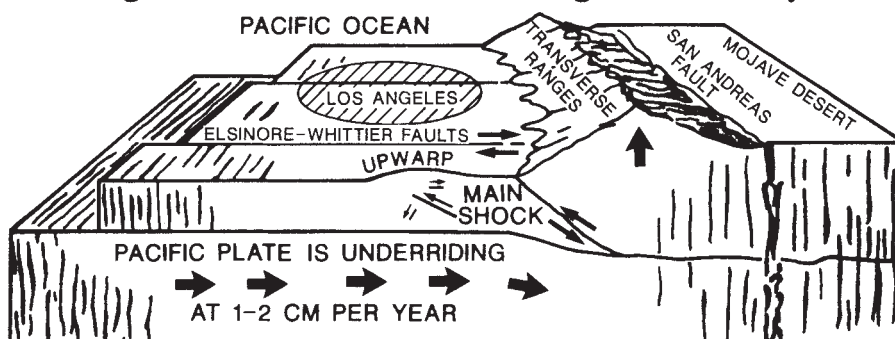
HUMAN immunodeficiency virus (HIV) has been isolated for the first time from patients in India with AIDS. Dr Pradeep Seth, a microbiologist at the All India Institute of Medical Sciences in New Delhi says it is not possible to say at present whether the virus is HIV-1, HIV-2, or a new strain.

The virus was isolated from the lymphocytes of three asymptomatic prostitutes by co-cultivation with phytohaemagglutinin-stimulated normal human lymphocytes. The presence of the virus was confirmed by indirect immunofluorescence using mouse monoclonal antibodies against HIV antigens p24 and p17, and by a reverse transcriptase assay with the culture supernatants.

Because of a lack of virus typing and containment facilities, the Indian isolates have not been further characterized. "Until this is done we are not sure if it is one of the known viruses or a new strain of HIV", says Seth. Typing and characterization of Indian HIV has acquired a new urgency in the light of the peculiar HIV-antibody distribution in the country. Although there are 145 known cases of positive antibody infection, only one full-blown AIDS case has surfaced.

K.S.J.

Geological basis of the Los Angeles fault system



Large earthquakes may hit Los Angeles when blocks of crust slip past one another at the San Andreas and Elsinore-Whittier faults. But, as the recent earthquake shows, trouble can also come when the underlying motion of the Pacific plate causes the crust above to shorten and crumple. Data on underlying folds produced by plate movement are now being sought from oil companies, which have had an interest in them for the oil they may contain. In the long run, moderate-sized earthquakes at faults that accommodate shortening may cause more damage than the occasional big earthquake from the San Andreas and Elsinore-Whittier faults. (Figure: Bob Wallace, USGS.)

Next-generation synchrotron to be built in Berlin?

West Berlin

DETERMINED not to fall behind in the race to build the next generation of synchrotron radiation sources, the West German government seems likely to approve a new 1.5 GeV storage ring in West Berlin. BESSY II (for Berliner Elektronenspeicherring-Gesellschaft für Synchrotronstrahlung) will have eight 'wigglers', essentially regions of corrugated magnetic field, to concentrate photon production in particular parts of the ring. This technique is also being used in synchrotrons in the same energy range in Italy, France and the United States.

BESSY II would build on its predecessor, BESSY I, a vacuum storage ring operating in West Berlin since 1982. By building BESSY II, West Germany could stay abreast of the Italian synchrotron planned for Trieste and the Advanced Light Source at Berkeley in the United States. None of these synchrotrons will approach the particle energies (up to 5 GeV) achieved by the European Synchrotron Radiation Facility under construction at Grenoble. BESSY II will cost roughly DM 80 million and the storage ring could go on-line by 1991.

Once built, BESSY II would be turned

over to a private corporation, following the precedent of BESSY I, which is jointly owned by the Max Planck Society, the Fraunhofer Society, the Hahn-Meitner Institute of West Berlin, DESY (the Deutsches Elektron Synchrotron) as well as by Siemens, AEG, Philips and Eurosil.

The location of a successor to BESSY I will depend in part on which way the political winds are blowing. West Berlin depends on West Germany for large-scale financial support and there has lately been talk that the city might lose some of its DM 800 million subsidy when the West German tax reform package comes into effect in 1990. A decision is expected by autumn 1988.

Support from potential users of BESSY II may help to tip the scales in its favour. According to Heinz Lindenberger, director of the Hahn-Meitner Institute, "the customers are queuing up already". A group at the University of Göttingen would use the new synchrotron to do real-time X-ray microscopy. The Nuclear Research Institute at Jülich would perform spin-polarized photoemission spectroscopy, and a group from the Fritz Haber Institute is interested in studying catalysis.

Steven Dickman

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Ready for BESSY II — the Berlin home for the new synchrotron.

UK compensates AIDS haemophiliacs

London

HAEMOPHILIACS who have become infected with the AIDS virus through contaminated blood products prescribed on the National Health Service are to benefit from a £10 million compensation fund to be announced by Mr John Moore, the Social Services Secretary, this week.

The biggest payments will go to the 21 haemophiliacs already suffering from the full-blown disease. Families of the 45 who have already died will also be compensated. There are currently 1,200 infected haemophiliacs in Britain.

The government is expected to insist that the fund does not represent compensation to avoid setting a precedent for special benefit schemes. A spokesman for the

Department of Health and Social Security (DHSS) said: "It is not compensation. The Haemophilia Society will have the money and will deal with it themselves". They are also unwilling to do anything that could be taken as an admission that supplies of the blood-clotting agent, factor VIII, were not safeguarded against contamination when the threat of AIDS was identified 4 years ago.

Last month the DHSS said that if haemophiliacs wanted compensation they would have to sue for negligence. Some patients are planning such action. But Moore appears to have convinced the government that the problems of haemophiliacs suffering from AIDS are exceptional.

Sarah Hargreaves

Lysenkoism still holding sway

London

THE Soviet popular-science monthly, *Priroda* (Nature) last month devoted an entire issue to the life and works of Nikolai Vavilov, in celebration of the centenary of his birth. Although, in the course of its 75 years' existence, *Priroda* has frequently celebrated the anniversaries of significant scientists and events, a single-theme issue of this kind is unusual — so all the more significant.

Vavilov was not only one of the Soviet Union's leading biologists and geneticists of the 1920s and 1930s, but also the most eminent victim of lysenkoism. Although he has long since been posthumously rehabilitated, this special issue has given *Priroda* an excellent opportunity to exhibit *glasnost* (openness) by revealing such hitherto unknown events and material as Vavilov's campaign to protect his genetics institute in 1938 and the drafts of research programmes and plans that he was prevented, by political pressures, from carrying out.

But the commemorative issue contains a warning that the fight against lysenkoism is not yet over. Among the book reviews is an attack on a series of pamphlets aimed, the reviewers claim, at rehabilitating if not Lysenko himself, then at least some of the leading figures of the Lysenko period, including, first and foremost, Michurin. One of these brochures "Michurin and Contemporary Biology", the reviewers note, actually credits Michurin with the formulation of five scientific laws, although Michurin (primarily a practical man) never made any such claim. The authors of the brochures, B.G. Ioganzen of Tomsk State University and E.D. Logachev, of Lemerovo State Medical Institute, have, said the reviewers, taken advantage of *glasnost* to issue works which are "dangerous" and liable to "disorientate the insufficiently informed reader", and to lead to opinions which are both philosophically unsound and harmful for agricultural and medical practice". They also attempt to discredit genetics by suggesting that it can lead to such unacceptable consequences as "social-darwinism, eugenics and racism". In the face of such "distortions" and "sophistry", the reviewers claim in their title, "we cannot keep silent".

It is "not by chance", the reviewers stress, that they have placed their warning in the Vavilov memorial number of *Priroda*. The prestige of the journal and the occasion, they hope, will help alert the Soviet ministries of Higher Education and of Health, as well as "wide circles of the scientific community" to the danger of a lysenkoist revival.

Vera Rich

Japan puts down the welcome mat for foreign researchers

Tokyo

UNDER fire for failing to make its contribution to the world store of basic research knowledge (but see below), Japan is struggling to remedy the situation. The issue has come to the fore in negotiations over the US-Japan Science and Technology Agreement (*Nature* 329, 662; 1987).

The impression of a one-way trade in researchers between Japan and the United States is confirmed by a Nomura Research Institute survey carried out in 1983 at the request of the Science and Technology Agency and by recent figures from the Japanese immigration department. Some 16,000 Japanese scientists and engineers visit the United States each year while only about 2,000 travel in the other direction.

The Ministry of Education, Science and Culture is trying to change the situation. It has applied for ¥373 million (\$2.7 million) in fiscal year 1988 to invite 100 postdoctoral fellows for a year from the United States and Europe, in addition to the 30 already invited from Britain, France and West Germany. The Science and Technology Agency has also applied for extra money for foreigners: ¥600 million (\$4.4 million) in fiscal 1988. Also, the Ministry of International Trade and Industry (MITI) is considering accepting US engineering postdoctoral fellows. As a result, Alexander DeAngelis of the US National Science Foundation office in Tokyo, says there will be a "substantial increase" in the number of American scientists visiting Japan in the next year.

But will postdoctoral researchers want to come? The ten fellowships available to British scientists each year are often undersubscribed and West Germany barely fills its quota. This unwillingness probably stems from the lack of employment opportunities for foreign academics, concern over the language problems and fear of isolation from the Western research community. Japan's laws were rewritten in 1983 to allow appointment of foreigners in faculty positions. But so far only 61 have been appointed in Japan's 95 national universities, and only a handful enjoy full tenure like their Japanese colleagues. Many of the appointees are Chinese, Koreans and Japanese-Americans.

Some new efforts are being made to employ foreign researchers at Tokyo University. The university has introduced a scheme that allows private companies to endow funds for guest and assistant professorships for periods of one to five years. Four companies (NEC, Nippon Telegraph and Telephone, CSK — a computer manufacturer — and Shin Nippon Steel) have endowed four posts.

The situation in government laboratories is even worse. There is only one foreigner on the staff of a government institute. When Dr Thomas Barnes of Britain joined MITI's Electrotechnical Laboratory in Tsukuba in September, it made the national news.

But it is the imbalance in the private sector that most concerns the United States. A large proportion of the researchers visiting the United States are from private companies. Many of the superconductor researchers at NEC and Toshiba were trained in US laboratories, as have many researchers in the Japanese biotechnology industry.

At the heart of the problem may lie

structural differences between the US and Japanese research systems. The bulk of Japanese research is funded by the private sector and is largely applications-oriented. Much of Japan's technology is the envy of the world and several foreign companies, such as Du Pont and ICI, have recently set up laboratories here to tap Japan's excellence. On the other hand, Japan's comparative weakness in basic research means that it has less to offer researchers at its universities and institutes.

Even gaining access to information on Japanese science and technology is not easy for US researchers. Japanese researchers can plug their computers into US databases: 8,000 Japanese users are already linked to the US DIALOG database. But Japanese databases and computer networks are not so advanced and much of the information is in Japanese.

David Swinbanks & Harry Glick

Japan's international standing improved by paper boom?

Tokyo

Is Japan's performance in basic research really as bad as critics think? A new report from the Ministry of Education, Science and Culture presents a more positive picture: in terms of number of publications, Japan now ranks second only to the United States in many fields of science.

The ministry analysed the country of origin of publications listed in four databases, INSPEC (published by Britain's Institute of Electrical Engineers), EMBASE (The Netherlands' *Excerpta Medica*), COMPENDEX (US Engineering Information) and *Chemical Abstracts* (US). The survey shows a remarkable increase in the number of Japanese publications over the past decade. In genetics, for example, contributions have shot up fourfold, and in 1985 surpassed those of

Britain, placing Japan second only to the United States. Japan also ranked second in 12 fields, including cancer research, electronics and electrical engineering and agricultural chemistry, and third in the other 11 fields analysed.

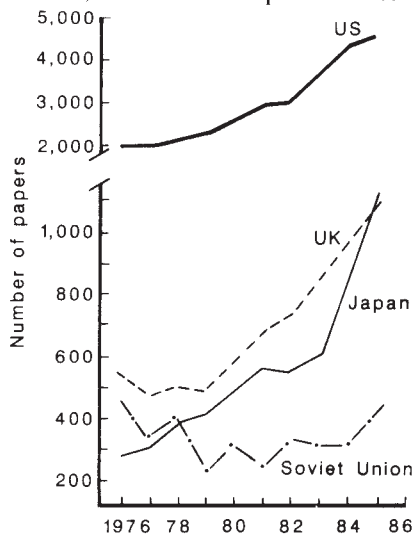
On release of the results last week, Japanese newspapers boldly declared that Japan is now number two in producing publications in leading journals, such as *Nature* and *Science*. And Hiroshi Ueki of the ministry's science and international affairs bureau says the survey shows that the quality of Japanese science is close to the best international standards.

But the ministry's survey drew data from more than 10,000 journals published in English, and the quality of many is open to question. For example, *The Japanese Journal of Applied Physics* announced in March that it would abandon peer review of papers on high-temperature superconductors because of a flood of contributions. And for many other journals, refereeing is more a formality than a strict assessment.

A different picture comes from a survey carried out by the Science and Technology Agency in the summer. By analysing the frequency of citations of Japanese papers, it ranked Japan fifth in basic science, well behind the United States, West Germany, France and Britain.

But can such measures be used to assess the standard of Japanese science? Many Japanese researchers complain that their work often goes uncited by Western researchers. It is well known that scientists tend to refer to the work of people they know, and as long as Japan remains comparatively isolated, its research will be underrated.

David Swinbanks



Comparison of academic publications in genetics, 1976-85.

Potential carcinogens are under-regulated says OTA report

Washington

A KIND of paralysis seems to strike federal agencies when it comes to regulating carcinogens. According to a new report* by the congressional Office of Technology Assessment (OTA), regulatory agencies have set standards for fewer than half the chemicals listed in the government's *Annual Report on Carcinogens*.

Although Washington's bureaucratic jungle has always made setting regulatory standards difficult, the report identifies new hurdles added by the Reagan administration that have further slowed the process.

Since 1978, the National Toxicology Program has been collecting data on potential carcinogens. So far, it has tested more than 300 compounds, and on the basis of interagency discussions concludes that there is sufficient evidence that some 148 of these should be listed as carcinogens in its annual report.

A staggering alphabet soup of federal agencies is responsible for regulating these compounds. The Environmental Protection Agency (EPA), the Food and Drug Administration (FDA), the Occupational Safety and Health Administration (OSHA), the Mine Safety and Health Administration (MSHA), the Consumer Product Safety Commission (CPSC) and the Office of Management and Budget (OMB) all have a role in establishing safety guidelines for exposure to carcinogenic substances. Overlapping jurisdictions are not uncommon. EPA may set environmental standards for a particular substance for emission into the atmosphere, but OSHA may have different standards for acute exposure in the workplace.

Given the potential for bureaucratic tangles, it is of little surprise that OTA found "apparent gaps in regulatory coverage". As an example, OSHA has determined that it has regulatory responsibility for 110 carcinogens listed in the report, but has set standards for only 17.

Less than one month after he took office, President Reagan signed an executive order requiring OMB to review all new regulations, and stating that "regulatory action shall not be taken unless the potential benefits to society for the regulation outweigh the potential costs". The idea was to relieve industry from needless interference from Washington, but critics have argued that not only does this slow the regulatory process, but it is inappropriate for OMB to usurp the authority of other federal agencies.

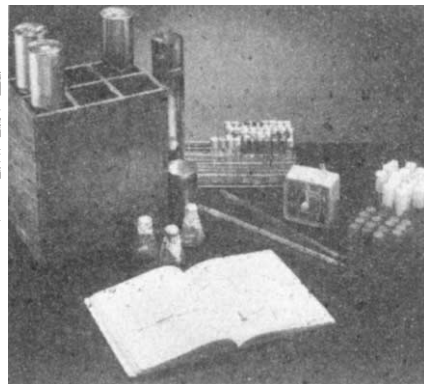
Karl Kronebusch, primary author of the

OTA report, says agencies do not feel they have neglected their regulatory responsibilities. He says they argue that the law requires them to evaluate benefits and risks of using chemicals that may be carcinogens. The cost of regulating must also enter into the agency's decision on whether to act. But Kronebusch says it is often hard to tell just what an agency has in mind when it declines to publish regulations on a particular compound, with many compounds languishing in regulatory limbo for years.

The OTA report admits that most cases of cancer are not caused by exposure to environmental carcinogens created by humans. But the report states, "those carcinogenic chemicals that can be identified specifically and can be controlled are important for those very reasons: they are avoidable."

Joseph Palca

Historic exhibits



ALTHOUGH the first plasmid was spliced artificially only 15 years ago, the Smithsonian Institution has decided that an historical exhibit on genetic engineering techniques is in order. The Smithsonian has created an exhibit, "The Search for Life: Genetic Technology in the 20th Century" that opens this week at the National Museum of American History in Washington, DC. The exhibit includes artefacts from the laboratories of researchers who contributed to breakthroughs in the understanding of DNA: Hershey and Chase's Waring blender, Bruce Merrifield's prototype peptide synthesizer, two base plates from the original Watson and Crick DNA model, and the laboratory notebook in which Stanley Cohen first outlined gene-splicing (pictured above). The narrated exhibit is intended to explain the significance of the new biotechnology to the layperson, and also includes a 'cell theatre' with screens that close around the viewer like the petals of a flower to simulate being in the interior of a cell.

C.E.

TPA freed for US use at last

Washington

THE long-awaited approval of Genentech's blood-clot-dissolving drug tissue plasminogen activator (TPA) was announced last week by the US Food and Drug Administration (FDA). As the first product of biotechnology with a large market — up to \$1 million million per year — TPA is expected to become the first 'blockbuster' of the biotechnology industry. But with the recent approval of streptokinase for intravenous use and future competition, Genentech may have to share the wealth.

Genentech plans to charge between \$2,000 and \$2,500 for one dose of TPA, under the trade name Activase, according to a company spokesperson. This makes it more than ten times as expensive as streptokinase. But Genentech believes that data suggesting that TPA may be twice as effective in breaking down clots will persuade physicians and patients to spend the extra money.

Genentech's strongest competition will come from 'second generation' versions of TPA made by protein engineering. Genetics Institute — the current leader in the race for the next generation product — has a version of the TPA molecule with an extended half-life in the bloodstream that the company hopes will make the drug more effective. Others working on the natural TPA molecule are Smith-Kline, Integrated Genetics, Eli Lilly and G.D. Searle. Genentech also has a second-generation TPA under development.

The FDA has been under heavy pressure to approve TPA since its decision in May to send Genentech back to the clinic to collect more data on the drug's long-term effect on reducing mortality (see *Nature* 327, 450; 1987). At that time, the FDA recommended against approval, citing the drug's infrequent side effect of inducing cranial bleeding and the lack of proof that clot dissolution decrease mortality from heart attacks. The decision drew strong criticism from cardiologists, patients and the media, who accused the FDA of allowing an internal battle to keep heart attack victims from receiving a drug that could save their lives. The FDA said the delay was due to Genentech's shoddy design of clinical studies.

After Genentech filed its additional data on 29 September, FDA took just seven weeks to give TPA final marketing approval. The company plans to begin shipping the drug to hospitals within 2–3 weeks. Genentech's foreign licensee, Boehringer Ingelheim, already markets TPA in Austria, West Germany, France, New Zealand, South Korea and the Philippines.

Carol Ezzell

*Identifying and regulating carcinogens. US Congress, Office of Technology Assessment, OTA-BP-H-42. US Government Printing Office, Washington, DC, 1987.

Foundation net to be widened for future prizes

Bern

SUSTAINED by the fortune of an anti-fascist publisher who fled from Italy in 1933, the Balzan Foundation is doing what it can to secure recognition for its prizes, awarded annually since 1978, as complements to the Nobel prizes. At the awards ceremony here on 13 November, officials of the foundation made it clear that they wish to cast their net wider than Italy and Switzerland, where the Balzan prizes have so far attracted most attention.

This year's prizes (see *Nature* 329, 279; 1987) went to a physical anthropologist, a human psychologist and a mediaeval historian. Each recipient will receive 250,000 Swiss francs (roughly \$180,000) to use as he likes. The prizes are also intended to help recipients "enrich their students".

The prizes are given specifically in areas not covered by the Nobel prizes and which are identified in advance. The 1988 prizes, which will be awarded in Rome, will be in the fields of applied botany, sociology and comparative literature. The value of the prizes will increase in 1988 to SFr300,000. Every third year, there will be a peace prize.

South African physical anthropologist Phillip Tobias, 62, head of the Department of Anatomy at the University of the Witwatersrand, was selected for his contributions to the study of early hominids. Tobias intends to use his prize for "some type of educational foundation to provide facilities for underprivileged students in South Africa".

Jerome Bruner, 72, divides his time between Harvard University and the New School for Social Research in New York.

A wealth of prizes awarded this week. Top: Drs Jerome S. Bruner, Richard Southern and Phillip V. Tobias, winners of the Balzan Foundation awards. Bottom: Professors Morris Cohen and Jan Hendrik Oort, recipients of the Kyoto prizes and Gerald F. Tape, winner of one of the Fermi awards.



Bruner is known as a cognitive psychologist. He was an early opponent of behaviourism. He is among those psychologists who assert that the development of a child reflects the culture and values with which he or she is surrounded. In a celebrated experiment in the early 1960s, he asked children from poor and rich families to judge the size of a dollar coin; the poorer children judged it to be larger.

British mediaeval historian Sir Richard Southern, of Oxford, the third recipient, has distinguished himself in his easy-to-read yet widely respected accounts of the religious and intellectual development of Europe in the Middle Ages.

Balzan, a publishing magnate, died in Switzerland in 1953 without leaving a will. The foundation was established in 1957 in settlement of a dispute between Italian and Swiss authorities, who disagreed about whether the Balzan fortune should revert to Italy upon the death of Angela Balzan, Eugenio's daughter and sole heiress. This is why the organization is actually a dual one, with the prizewinners selected by the Milan branch and the money managed by the Zurich branch.

Steven Dickman

Prizegiving season arrives

Tokyo

It is the season for Japan to give away big money prizes to scientists. Last week it was the Kyoto prize and Honda prize, next week the International prize for Biology.

Professor Jan Hendrik Oort, 87, of the Netherlands, and Professor Morris Cohen, 75, of the United States, are the two scientists among this year's three winners of the Kyoto prize. The prize was set up in 1985 by Kazuo Inamori, founder of Kyocera Corporation, a fast-growing ceramics manufacturer. Recipients must be people who "are sensitive to their own human fallibility, and who have a deeply rooted reverence for the universal spirit". Oort's reverence for the Universe has no doubt increased — he gets his ¥45 million (\$330,000) prize for elucidating the structure and dynamics of the Galaxy — while Cohen is rewarded for developing new materials, like ultra-high strength steels, shape-memory alloys and ceramics.

The Honda prize, awarded by a foundation set up by Honda Motor Company, is given to individuals or organizations that promote "ecotechnology". According to the foundation, this means technology that "harmonizes the overall environment of human activity". This year's award of ¥10 million (\$75,000) goes to Professor Jean Dausset, 71, of the Collège de France in Paris. Dausset received the Nobel prize for Medicine in 1980 and has promoted a worldwide group called the "Universal Movement of Scientific Responsibility".

The International prize for Biology was established in 1985 to celebrate the sixtieth anniversary of the Emperor's accession to the throne, and is intended to become Japan's "Nobel prize" of biology. The ¥10 million prize goes to Professor John Bertrand Gurdon of Cambridge University, who is best known for his demonstration in the 1970s that it is possible to clone frog (*Xenopus*) eggs.

David Swinbanks

Fermi awards for atomic energy advances

San Francisco

THE US Department of Energy (DoE) has chosen a researcher and an administrator to receive the 1987 Enrico Fermi Awards. Luis W. Alvarez and Gerald F. Tape will each receive \$100,000, a gold medal and a presidential citation.

The Fermi awards have been given each year since 1954 to honour scientific and technical achievement in the development, use or control of atomic energy.

Tape, 72, is past president of Associated Universities Inc., a non-profit organization that manages Brookhaven National Laboratory and the National Radioastronomy Observatory. His award acknowledges his contributions to the advancement of nuclear power and the non-proliferation of nuclear weapons.

Tape served as the US representative to

the International Atomic Energy Agency from 1973 to 1977, and was instrumental in the implementation of the nuclear non-proliferation treaty.

Alvarez, 76, is professor emeritus of physics at the University of California at Berkeley and senior scientist at Lawrence Berkeley Laboratory. He received the 1968 Nobel prize in physics for his work on the development of the hydrogen bubble chamber and subsequent use of the bubble chamber to reveal previously unobserved, shortlived classes of subatomic particles.

The advance revolutionized the field of high-energy physics and eventually led to the quark model, the currently accepted description of sub-atomic matter.

The awards will be presented at a ceremony in Washington in December.

Marcia Barinaga

Battles on radiation safety

SIR—Your recent leading article ("Hard battles on radiation safety" *Nature* 329, 185; 1987) is a generally accurate and evenhanded statement, but it does not seem adequately to stress the magnitude or the relative importance of the problems involved.

As you say, the "dose" limits (for general populations) recommended by the International Committee on Radiological Protection (ICRP) amount to less than the doses received from background (radiation). As background is subject to substantial geographic variation, one might assume that comparisons of the incidence of deleterious effects (especially cancer mortality) in regions where the background levels differ could permit an estimate of the risks posed by ICRP recommendations. A number of such studies (often involving vast populations) have been made and more often than not they have indicated that the cancer mortality is less where the background level is higher. The value of the results of any of these projects is, however, dubious, as evidenced by quite substantial variations in cancer mortality among regions having similar background levels. This rather unsurprising finding would seem to doom any epidemiological evaluation of the risks implied by ICRP recommendations.

Further analysis of the data from Hiroshima and Nagasaki, and possibly also from Chernobyl, may improve our imprecise estimates on the induction of cancer and maybe even of genetic damage when caused by substantial fractions of a sievert. I do not believe that anyone with knowledge of the epidemiological difficulties involved expects any meaningful evidence at 1 or even 10 mSv.

Such data can therefore be used only in conjunction with extrapolations. While casting some doubt on the notion of proportionality, you state that the existence of a threshold for cancer induction is "unlikely on theoretical grounds". The theory involved is quite naive, as shown by the well-established fact that the natural incidence of a few animal tumours is reduced by modest doses of radiation that cause excess incidence of other tumours. But in experimental radiobiology, too, our knowledge of what happens at a few mSv is generally sparse.

In view of these uncertainties, it cannot be said to be unreasonable for ICRP to employ linear extrapolations in attempting to quantify radiation risks. It has based such estimates on the best available data, and it will no doubt revise them, if warranted, when seemingly better data become available. What should be criticized is its hesitancy to state clearly that the risk figures are nominal rather than actual.

The significance of radiation risks is, however, largely a polemical issue. Most people (and probably many of the signatories of the petition to the ICRP) would not hold the hazards of background radiation worth considering when deciding whether to move from Florida to Colorado (or the other way). It is difficult to discount the impression that the petitioners were motivated by politics or ideology rather than by genuine concern over radiation safety.

HARALD H. ROSSI

*Department of Radiation Oncology,
College of Physicians & Surgeons
of Columbia University,
PO Box 21,
Irvington,
New York 10533, USA*

Emergency plans

SIR—The analysis in "US nuclear elephants stay white" (*Nature* 328, 561; 1987) is seriously flawed by the author's lack of understanding of how emergency planning is accomplished under the federal system of the United States. Under that system, no agency in the federal government, including the Nuclear Regulatory Commission (NRC), is responsible for developing an emergency plan for a nuclear reactor. Rather, the development of such a plan is the responsibility of the state and/or local governments in whose jurisdiction the particular reactor is located.

Thus the acceptability of an emergency plan is not simply a question of whether the local authorities think it makes sense: if those local authorities decide not to participate in that plan, there is no workable plan. Even the NRC acknowledges this in its assumption that a utility plan is a workable substitute only because, in the event of an emergency, state and local governments will be compelled to participate despite previous disclaimers of any intention to do so.

These realities, and the confusion over evacuation authority after the accident at Three Mile Island, left the NRC no practical choice in 1980 but to institutionalize the critical role of local interests in emergency planning. The issue, therefore, is not whether the NRC made a mistake in 1980 by giving local interests such a powerful voice in national energy interests. Rather, it is why local interests at only two of over sixty sites have chosen not to participate in emergency planning.

One answer, which seems to have been dismissed too cavalierly, is economics. The traditional rate basing of Shoreham would further exaggerate Long Island's unenviable position of already having one of the nation's highest electric rates. It is

this local economic reality that seems to have led local Long Island politicians to seize on emergency planning as a way to stop the licensing of Shoreham until a workable rate treatment could be achieved. But now, having mounted that proverbial tiger's back, no one can dismount gracefully. That will happen only when the affected interests are candid about their concerns.

By the way, Peach Bottom is in Pennsylvania.

SHELDON L. TRUBATCH

*5825 S Blackstone,
Chicago, Illinois 60637, USA*

Staying on

SIR—In response to the letter of J.G. Wilson (*Nature* 328, 288; 1987), the fundamental absurdity of racial discrimination is that it is based on something beyond the control of the person(s) at which it is aimed, namely being born into one race rather than another. Wilson is guilty of exactly the same offence. He condemns all (white) South African scientists (who "choose to enjoy the advantages of living comfortably in South Africa, at the expense of the black population" — my italics) to be "denied the benefits of attendance at international conferences and so on".

To condemn South African scientists as racists, to be unaware that much of what politicians do is quite unacceptable but beyond the control of scientists and to hold it against us that we were born white and in South Africa is naive and just another step in a whirlpool of racism. I find it sad that scientists, who often claim to be the most rational of academics, have to resort to racism in their attempts to fight the very same.

Does Wilson recommend that South African scientists all leave the country, become politicians or revolutionaries? Many scientists have already left this country but many of us are staying. We are staying not because we see ourselves in privileged positions or because we want to suppress black people, but because by continuing to do research we feel we are working towards a better life for all in this country, no matter who is in government. We are not unthinking dinosaurs but have evaluated the very real uncertainty of our future and have chosen to stay because we are Africans, and do not wish to be "re-settled" anywhere else on this planet as was arrogantly suggested by W.D. Stein (*Nature* 328, 374; 1987). We are able to accept the responsibility of our situation, even though many of us have not brought it about.

GEORGE DENNILL

*Plant Protection Research Institute,
Private Bag X5017,
Stellenbosch 7600,
South Africa*

Making quantum mechanics relativistic

A novel attempt to reconcile dull old Schrödinger's equation with special relativity would be more winsome if it had more explicitly acknowledged its purpose.

DIRAC'S achievement in the late 1920s notwithstanding, people seem still to be brooding about the best way of making quantum mechanics relativistic. That, at least, is the burden of a long article by A. Kyprianidis, from the Poincaré Institute in Paris, now published (*Physics Reports* **155**, 1; 1987). It is only appropriate that the issue should have been raised from an address in Paris; the French physicist Louis de Broglie was always slightly alienated from the mainstream in the 1920s and 1930s by his insistence that there is something odd about the treatment of time, in the Schrödinger equation and elsewhere in elementary quantum mechanics.

The problem is simply put. The time-dependent form of Schrödinger's wave equation is a linear differential equation for the wave function ψ in which the first derivative of ψ with respect to time is equated with the effect on ψ of an operator that includes second-order differentials with respect to the space coordinates. For a single particle, the space-derivative side of the equation represents the Hamiltonian of the particle, H , the sum of its kinetic energy and potential energy and that the time-derivative side of the equation represents the energy, E , which amounts simply to $H\psi = E\psi$.

The simple rule for arriving at the differential equation is that the kinetic energy part of the Hamiltonian should be written in terms of momentum $\mathbf{p}$, or $\mathbf{p}^2/2m$, where m is the mass, whereupon each of the three components of momentum must be replaced by the differential operator $\hbar/i \cdot d/dx$, where x is one of the three space coordinates, while E is replaced by $-\hbar/i \cdot d/dt$. (In all this, i is the square root of -1 .)

Plainly, the time and space coordinates are not dealt with on an equal footing, and so Schrödinger's equation cannot be relativistic. De Broglie was well within his rights to protest that time is dealt with as if it were simply a parameter describing the evolution of the system the equation represents, although his assertion that Schrödinger's equation says nothing of the mutual uncertainty in measurements of time and energy is more disputable.

Historically, there was no obvious way to generalize Schrödinger's equation. The essential difficulty is that special relativity relates the energy of a free particle to its momentum in a more complicated fashion than in classical mechanics. Indeed, the

energy of a free particle is $c\sqrt{(\mathbf{p}^2 + m^2c^2)}$, which makes for an awkward differential equation when appropriate partial differential operators are substituted for the components of the momenta.

Unfortunately, the obvious way round the difficulty, that of deriving a differential equation by writing $E^2 = c^2\mathbf{p}^2 + m^2c^4$ and substituting differential operators as above, leads to the Klein-Gordon equation, which suffers from several defects, not least that it will not allow of the simple interpretation that the square of the magnitude of ψ should represent the probability density, if only because this quantity may be negative.

Dirac's achievement was his demonstration that Schrödinger's equation can nevertheless be generalized by analogy into a form in which both the space and time derivatives, representing momenta and energy respectively, appear only to the first order. One penalty (but, of course, it is the prize) is that such an equation compels the existence of Pauli's electron spin as well as the electron states with negative energy which, in the sequel of Dirac's argument, led to the prediction that positrons should be holes in this limitless sea of negative electrons.

That seems part of the incentive for what Kyprianidis has attempted, although Dirac rates a mention only once (in the subtitle to de Broglie's book appearing among the references). The argument is nevertheless instructive, even in parts entertaining. The stated purpose is to stay with the ordinary quantum mechanics of the kind described by Schrödinger, but to do so relativistically, but to avoid using the language of quantum field theory.

In short, the objective is to be able to represent the states of a single particle by wave functions ψ , to represent stationary states of a single particle by particular wave functions (which must be function of the space and time coordinates) and to require, as in the familiar Schrödinger calculus, that these stationary states are all orthogonal to each other when integrated over all coordinates from $-\infty$ to $+\infty$.

The essential trick is the definition of an evolutionary parameter distinct both from the ordinary or 'fourth' coordinate and from the proper time of the particle. The progress of a particle through space and time is represented by a stochastic process (strictly, a Markov process) in which the state of affairs at one value of the evolutionary parameter is determined by the state of affairs at earlier and/or later

values (allowing for such happenings as the creation and the annihilation of the particle).

The result is a Schrödinger equation in which the wave functions have five coordinates (three representing space, one time and one evolutionary direction). Loyal to his fellow city-dweller de Broglie, Kyprianidis claims that the objections of de Broglie to the wave equation are overcome. He says that the "advantages of this formalism [over what?] are immense", but also acknowledges that "explicit solution of open problems in the frame of this formalism will be the best proof of its validity".

There are some points in the argument, however, when a suspension of disbelief is necessary. For example, to arrange for the orthogonality of the wave equations representing stationary states, it is necessary to require that they should vanish at infinity in all direction, in time as well as space. Innocents will find themselves asking just what physical significance there can be in the concept of a free and isolated particle whose wave function gives zero probability of existence on all sufficiently distant space-like surfaces, let alone in the use of such a set of functions, however orthogonal they may be, to represent more general states of existence.

Sadly, what starts off as an appealing argument becomes, before its end, more than a little contrived and, in the process, introspective, as if the author is writing principally for the people whom he knows will understand what he intends. To say this is not to complain that the mathematics becomes more complicated as the pages turn, but that there is a hidden agenda — to find a way off the awkward conceptual hooks of quantum mechanics.

The sense one has of being let down by articles of this kind derives from an easily identified deficiency — Kyprianidis, while referring to the Klein-Gordon equation as an earlier (but unsatisfactory) landmark on the road to relativistic quantum mechanics, fails to say what he thinks is wrong with Dirac's equation or, for that matter, with the development of field theory in the past forty years (although Feynmann does win a reference in his own right). Instead, his problem is presented to innocent readers as one that is about to be tackled properly for the first time. It is forgivable that authors should write like that, but should not editors ask that their readers should be given a little help?

John Maddox

Evolution of helicity

A cause for reflection?

John Galloway

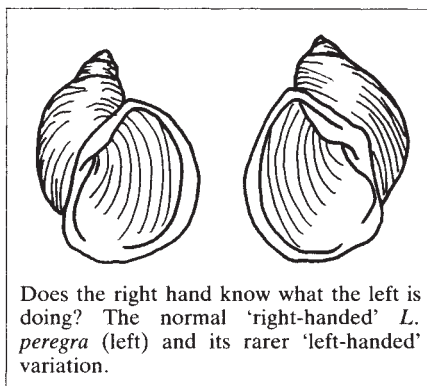
"It is no accident", Martin Gardner says¹, "that *Alice through the Looking Glass* is filled with references to mirror reversals and asymmetric objects. The helix itself is mentioned several times. Humpty Dumpty compares the toves to corkscrews . . .". Gardner goes on to detail plenty of other examples. The helix, because of its elementary geometry and distinctive appearance, is the clearest instance of an enantiomorphic object, one that, as Alice says, "goes the other way in a mirror". The helix and its mirror-image are identical in all respects except their screw sense. Although this distinction can be ignored from the standpoints of pure geometry and pure group theory², any helical structure is actually available as either or both hands. And the extent to which a helix in nature appears with both hands (or not), and why, provides one of the best puzzles in the science of form. A recent paper³ by M.S. Johnson looks at gastropods, which provide an excellent manifestation of this puzzle, and one which should open a window on the developmental processes underlying the contrast between left and right in living things.

The conchospirals of gastropod shells — which are true helices² — and the twisted animals they contain are almost, but not quite, always right-handed (dextral). (For completeness I should add that very occasionally shells contain the wrong hand of snail.) Most species are right-handed, although a few, like those of some land snails, such as Clausiliidae and some species of *Partula*, are left-handed (sinistral). Within species that are right-handed, a low frequency of left-handed conchospirals is sometimes found. The rarity of left-handed conchospirals has attracted the attention of natural historians as separated in time and philosophy as Aristotle, D'Arcy Thompson and Steven J. Gould⁴⁻⁶. The puzzle presented by this rarity has two aspects: what developmental mechanisms control hand; and what evolutionary mechanisms keep left-handedness so rare?

This strong tendency towards right-handedness tempered by a gesture towards left-handedness is perhaps a nice example of what G.K. Chesterton called "a secret treason in the Universe". It is sometimes, although by no means always, a feature of other helical forms. Vermeij reported⁷, for instance, that fossil conspiral nautiloid and ammonoid cephalopods are found with roughly equal frequencies of hand. Studies of the trochospirally coiled foraminifera seem to show that — for some of them at least — coil direction has

changed over time from being equally distributed between left and right to being predominantly single-handed.

Many types of bacteria are helical. Of the 17 identified species of the *Aquaspirilla*, twelve are right-handed and three are left (two are straight). It is claimed that within a species all individuals have the same hand. What about helical molecules? All naturally occurring α -helix is right-handed with, as far as I know, the single exception of a short length of left-hand helix in thermolysin. Indeed, all



structural features of proteins where there is a choice of hand tend to favour one hand over the other.

Steric constraints apparently lie behind this molecular partiality and if these permit one hand, almost without exception, they forbid the other. Could steric considerations similarly explain the preference of right- over left-handed conchospirals? Perhaps surprisingly, the answer may well be yes — at least in part — surprisingly because the steric considerations in this instance are behavioural.

Can evolution discriminate against a left-handed conchospiral, compared with a right-handed one? If the two are exact mutual mirror-images, then on the face of it the answer seems to be no. It is almost inconceivable that the possession of a right-handed rather than a left-handed sense of twist could by itself confer any advantage on its possessor. Of course, left- and right-handed shells are not true mirror-images of one another — their shapes differ in what have been claimed to be fairly consistent ways. This being the case, sinistral rarity might be a consequence of disadvantages conferred by these other concomitant structural features⁸. The case for understanding evolution through the notion of a *Bauplan* — with its consideration of organisms as wholes — and against adaptationism has been rehearsed entertainingly by Gould and Lewontin⁹.

It does seem a pity, though, to have to discard such a distinctive and elegant discriminatory feature between left- and right-handed shells as their mirror dissymmetry — even if this is only approximate. Simplicity may be a dangerous principle and 'mistrust it' a good maxim, but it can be discarded too precipitately. And, indeed, Johnson now suggests³ that in *Partula*, shell shape and screw sense appear to be evolving independently. He suggests that what evolution is actually selecting against is the relative rarity of the left-handed snails, and points out that even more striking than the general rarity of sinistral forms is "the rarity for variation in handedness". Of the named species of *Partula* that are found on the islands of Tahiti and Moorea, eight are right-handed, three left and only two polymorphic for sense of twist. Even in these polymorphic species, populations tend to consist of only one hand.

How could evolution discriminate against a rare left-handed form in a predominantly right-handed population and *vice versa*? Sturtevant⁹ may have made the significant observation here. Commenting on rare left-handedness in the European pond snail *Lymnaea peregrina*, he pointed out that the left- and right-handed forms might not be able to mate with each other, not having a compatible screw sense! Handedness in snails is a simple mendelian characteristic (although with the interesting feature that it is an example of maternal inheritance — the hand of coiling is determined not by the individual snail's own genes but by those of its mother). Presumably an effect of the left- and right-handed forms not mating with each other would be to create separate breeding populations. An individual of the rarer form will have relative difficulty in finding a mate of the same hand, thus keeping the rare form rare or, in favourable conditions, forming geographically separated right- and left-handed populations.

Although this argument would not explain the origins of the predominance of right-handed forms, it would link the rarity of left-handed species with the rarity of left-handed individuals in predominantly right-handed populations, and suggest how the imbalance could be maintained. Johnson¹⁰ has shown that Sturtevant's supposition is essentially correct, snails of opposing hands mating relatively rarely; but he has also discovered that nature, not content with making mating difficult, contrives that the fruits of any unions that do take place are less numerous in the rarer than the commoner form. An evolutionary advantage for a snail to change its screw sense would be to prevent mating with a different but closely related form of the same species, an explanation first put forward in respect to the switch from left- to right-handedness in *P. suturalis*¹¹.

An evolutionary mechanism combining sex with elementary geometry has a lot going for it, but it still does not explain why the predominant sense of twist in nature is right-handed. It is difficult to see how this problem could be attacked except through the explicit developmental mechanism that controls hand, and here, it seems, simplicity must be discarded. Interesting but probably ultimately unhelpful observations for understanding gastropods have been made on the effect of temperature on the screw sense adopted by some other helical organisms. A considerable body of circumstantial evidence supports the idea that water temperature has determined the hand of some foraminifera tests. Left-handedness is associated with low, and right-handedness with higher temperatures, in *Globigerina truncatulinoides* and *G. pachyderma*, for example. A temperature-induced switch of hand in the highly organized multicellular structure of *Bacillus subtilis* 'macrofibres' has also been observed and studied in some detail (see, for example, ref. 12). But are either of these inversions (reflections) genetic in origin?

Better, then, to study the gastropods directly. An observation made very early on in *Lymnaea* may offer a starting point. A trait inherited maternally should be seen in all the offspring from the same hatching. Thus, any broods of snails should be exclusively left- or right-handed. In practice, some broods possess a few snails of the opposing hand, and in sinistral broods the mutation rate to dextrality is quite high¹³. Freeman and Lundelius have proposed¹⁴ a genetic model for this phenomenon. The power of their model — or alternatives to it — to explain predominant right-handedness needs exploring.

What is now known from studying *Lymnaea* is that the dextral gene is expressed during oogenesis and its product — whatever it is — controls the symmetry of the pattern of very early cleavage. An injection of cytoplasm from dextral eggs reverses the early cleavage pattern of sinistral eggs, but one from sinistral eggs does not influence dextral ones¹⁴. To resolve this apparent paradox, it is

suggested that the normal form of *L. peregra* is actually left-handed, switchable to right-handed by the protein product of the dextral gene. Freeman and Lundelius speculate that what they have found in *Lymnaea* would be found exactly reversed in a left-handed species like *P. suturalis*. It is difficult to imagine that this will not turn out to be so, and then equally difficult not to agree with the authors in concluding

that the path from gene locus to cleavage pattern is likely to be far from straightforward — so far that, rather like Humpty Dumpty looking for the hippopotamus, the best instrument to be carrying on the search might be a corkscrew. □

John Galloway is at the Medical Research Council, 20 Park Crescent, London W1N 4AL, UK. The views expressed here are the personal views of the author.

Oceanography

Organic matter in sea water

John I. Hedges

IN contrast to the impression given by the abundant plant and animal life of coral reefs and other productive marine environments, more than 99 per cent of the organic matter in the ocean is dead, dissolved and extremely dilute. The complex mixture of organic molecules comprising this solution of degraded biochemicals, however, has a great effect on many ocean processes and plays an active role in the global carbon cycle. On page 246 of this issue¹, P.M. Williams and E. Druffel present the first detailed analysis of the radiocarbon (¹⁴C) activity of dissolved organic matter in the water column of the open ocean. Their results provide new insights on the sources and dynamics of this key sea-water component.

There are three reservoirs of reactive organic matter in nature, each of which holds approximately as much carbon as occurs in atmospheric CO₂ and thus has the potential to affect the concentration of this climatically critical gas². Two of these carbon reservoirs occur on land in the form of living plants and degraded organic matter in soils; the third, organic matter dissolved in sea water, is by far the least understood.

Many of the uncertainties about dissolved organic matter in sea water stem from the fact that less than 30 per cent of the component molecules have been identified¹. This lack of detailed structural information makes it extremely difficult to identify biochemical precursors or to judge the probability of different chemical and physical transformations. Structural characterizations, however, are challenging because most analytical techniques for organic compounds are highly selective, so that many different methods are required comprehensively to characterize complex natural mixtures. In addition, few methods are sufficiently sensitive for direct quantification at the low concentrations (parts per million to parts per billion (10⁹)) found in sea water, and no technique has yet been demonstrated for representatively concentrating dissolved organic matter from the mass of sea salt, which is more than a

thousand times greater.

The most successful isolation method for dissolved organic matter to date, involving selective adsorption onto synthetic resins, typically recovers less than 10 per cent of the total material. This relatively nonpolar fraction of 'marine humic substances'^{3,4}, consists primarily of acid-rich aliphatic polymers with molecular weights of 500–1,000 that bear little resemblance to any known biochemical or humic substances in soil. Because it is largely uncharacterized, the total amount of organic matter dissolved in sea water can be quantified only by complete combustion to CO₂. This common procedure measures dissolved organic carbon (DOC), which is equivalent to about half the organic matter present.

There is growing evidence that, although only a trace sea-water component, DOC is involved in various marine processes^{5,6}. Being surface-active, for example, DOC concentrates at the ocean surface where it can undergo photochemical reactions as well as affect the transfer of light, gases and wind energy. This tendency to concentrate at interfaces also causes DOC to adsorb onto particles suspended in sea water, thereby changing their charge and surface properties. Finally, some DOC components chemically combine with dissolved metals to form complexes that affect the chemical properties of sea water and phytoplankton production⁶.

In the light of these many important roles and the dearth of complementary information, the dissolved-organic-carbon profile reported in this issue is significant in at least two ways. First, the almost parallel exponential decreases in the radiocarbon content of both dissolved inorganic and organic carbon within the upper kilometre of the central North Pacific (Fig. 1a of ref. 1, on page 247) suggest that both features have a common origin in circulation processes operating in the upper ocean on a timescale shorter than the thousand-year mixing time of the global ocean. This finding provides the

1. Gardner, M. *The Wasp in a Wig* (Lewis Carroll Society of North America and Macmillan, London, 1977).
2. Coxeter, H.S.M. *Introduction to Geometry* (Wiley, Chichester, 1969).
3. Johnson, M.S. *Evolution* **41**, 672 (1987).
4. Gould, S.J., Young, N.D. & Kasson, W. *Evolution* **39**, 1364 (1985).
5. D'Arcy Thompson, W. *On Growth and Form* (Cambridge University Press, 1917).
6. Gould, S.J. *Nat. Hist.* **94**, 17 (1985).
7. Vermeij, G.J. *Nature* **254**, 419 (1975).
8. Gould, S.J. & Lewontin, R.C. *Proc. R. Soc. B* **205**, 581 (1979).
9. Sturtevant, A.H. *Science* **58**, 269 (1923).
10. Johnson, M.S. *Heredity* **49**, 145 (1982).
11. Clarke, B. & Murray, J. *Biol. J. Linn. Soc.* **1**, 31 (1969).
12. Favre, D., Thwaite, J.J. & Mendelson, N.H. *J. Bact.* **164**, 1136 (1985).
13. Boycott, A.E. & Diver, C. *Proc. R. Soc. B* **95**, 207 (1923).
14. Freeman, G. & Lundelius, J.W. *Wilhelm Roux's Arch. dev. Biol.* **191**, 69 (1982).

most definitive evidence to date that recently formed, dissolved organic molecules are being mixed rapidly into the upper ocean. This means that labile DOC may be an important source of food deeper in the ocean than was previously thought, and that synthetic organic compounds also could penetrate rapidly and deeply into the sea.

Second, DOC in the deep ocean has a mean age of roughly 6,000 years (approximately twice the previously suggested value⁷), which shows how unreactive this material is and refines significantly the global carbon cycle. This more precise radiocarbon measurement indicates that organic molecules dissolved in the deep sea on average survive approximately six exposures to concentrated light and life at the ocean surface during global circulation and are even more refractory than previously indicated. In addition, it is now clear that the amount of carbon that must enter or leave the ocean DOC pool per year (10^{14} grams of carbon; ref. 1) is much less than 1 per cent of the organic matter synthesized annually by marine plankton and only half the mass of DOC discharged into the ocean annually by rivers.

This result, together with the estimate that the amount of organic matter deposited in marine sediments does not exceed a quarter of the DOC input from rivers, indicates that most of the DOC entering the ocean from land, and most of it cycling within the ocean itself, must be destroyed by oxidation to CO_2 . This implication, supported by the observation that the organic materials in all three reservoirs are compositionally dissimilar^{4,7-10}, is surprising because both riverine and deep-ocean DOC are relatively resistant to biodegradation¹⁰. Ironically, Williams and Druffel's measurements¹ not only indicate that deep-ocean DOC is more refractory than previously thought, but also point towards an as-yet undiscovered process by which the sea eventually recycles this huge reservoir of recalcitrant organic molecules back to inorganic carbon. □

1. Williams, P.M. & Druffel, E.R.M. *Nature* **329**, 246-248 (1987).
2. Trabalka, J.R. *Atmospheric Carbon Dioxide and the Global Carbon Cycle* (US Department of Energy, 1985).
3. Stuerner, D.H. & Harver, G.H. *Nature* **250**, 480-481 (1976).
4. Meyers-Schulte, K.J. & Hedges, J.I. *Nature* **321**, 61-63 (1986).
5. Williams, P.J. leB. *Chem. Oceanogr.* **2**, 301-363 (1975).
6. Goldberg, E. & Bard, A. *Modern Chemistry and Chemical Technology Applied to the Ocean and Its Resources* (Pergamon, Oxford, in the press.)
7. Williams, P.M., Oeschger, H. & Kinney, P. *Nature* **224**, 256-258 (1969).
8. Romankevich, E.A. *Geochemistry of Organic Matter in the Ocean* (Springer, Berlin, New York, 1984).
9. Williams, P.M. & Gordon, L. *Deep Sea Res.* **17**, 19-27 (1970).
10. Thurman, E.M. *Organic Geochemistry of Natural Waters* (Junk, The Hague, 1985).

John I. Hedges is in the Department of Oceanography, University of Washington, Seattle, Washington 98195, USA.

Virology

Vector-directed interleukin expression and infection

Brigitte A. Askonas

THERE is a continued need to develop vaccines against serous infections that do not permit the use of attenuated infective particles. Purified components of pathogens are often poorly immunogenic and in most cases do not induce relevant responses of cytotoxic T cells. One approach to vaccine development is to insert genes encoding single viral antigens into vaccinia virus^{1,2}, and an interesting novel variant of this method has now been developed by Ramshaw and colleagues, who reported their results in *Nature* last month³, and by Flexner *et al.*, who report their results on page 259 of this issue⁴.

In the standard approach^{1,2}, the inserted gene as well as vaccinia virus (VV) proteins are expressed by the infected cells *in vivo* or *in vitro* to mimic antigen presentation in a normal virus infection. The recombinant VV proteins are good inducers of antibodies and different T-cell responses, and are excellent tools to study the T-cell repertoire for viral components (for example, influenza or respiratory syncytial virus components; refs 5, 6). But vaccination using such constructs can lead to disseminated vaccinia infections in immunosuppressed hosts. In their new work, Ramshaw *et al.* and Flexner *et al.* insert the gene encoding human or murine interleukin-2 (IL-2) into vaccinia virus that can contain additional viral genes such as influenza haemagglutinin (HA) or nucleoprotein (NP). IL-2 amplifies T-cell responses and also acts on B cells of the immune system. Ramshaw *et al.*³, using athymic nude mice as a model of an immunodeficient host, demonstrate that infection (local or intravenous) of IL-2/HA/VV causes clearance of vaccinia virus by day 11, whereas HA/VV leads to persistent vaccinia infection. In normal mice, inclusion of the IL-2 gene in recombinant VV does not influence the rate of virus clearance.

Flexner *et al.* in this issue⁴ present more extensive data in a model of disseminated VV infection. Inbred athymic nude mice die within a few days following intraperitoneal infection with recombinant vaccinia virus constructs. Inclusion of the IL-2 gene in the vaccinia virus vector results in survival for more than 60 days. It is of particular interest that evidence for IL-2 production in host mice could be obtained even 6 days after infection. This raises the question as to which cell types are able to produce IL-2 following vaccination, because athymic mice have no functional T cells. It is possible that

T-precursor cells are activated to differentiate into lymphokine-activated killer cells, with dramatic results: although vaccinia infection is not totally prevented, VV titres in affected tissues are strikingly low.

In normal host mice of the BALB/c strain, infection by recombinant vaccinia virus with inserted influenza HA or NP genes results in complete or partial protection, respectively, against lethal challenge with the homologous influenza virus. Co-expression of IL-2 in these vectors does not enhance partial protection nor does it overcome lack of protection by NP/VV in mice of the B10.A(5R) strain. The latter strain shows no NP-specific cytotoxic T cells as it expresses only NP non-responder class I alleles of the major histocompatibility complex⁷.

The preliminary data of Flexner *et al.*⁴ show variable effects on antiviral antibody formation by the IL-2/HA/VV vector, depending on the mouse strain and route of infection. The immunological effects of including the IL-2 gene in the recombinant vaccinia virus need to be studied, and particularly the long-term priming of B- and T-memory cells, the main aim of vaccination to protect against exposure to natural infection.

The concept of using recombinant vaccinia virus for clinical applications remains questionable, but it should be possible to develop additional vectors for the expression of desired growth or differentiation factors for possible clinical use. The observations by Flexner *et al.*, that IL-2 is produced for several days *in vivo*, is encouraging, because IL-2 and many interleukins or interferon- γ have very short half-lives when administered directly *in vivo*. Thus, continued production of a growth factor *in vivo* may modulate responses in severe immune disorders. The next step is to analyse the immunological effects induced by co-expression of IL-2 in recombinant vaccinia in athymic or other immunodeficient or immunosuppressed hosts. □

1. Mackett, M., Smith, G.L. & Moss, B. *Proc. natn. Acad. Sci. U.S.A.* **79**, 7415-7419 (1982).
2. Moss, B. & Flexner, C. A. *Rev. Immun.* **5**, 305-324 (1987).
3. Ramshaw, I.A., Andrew, M.E., Phillips, S.M., Boyle, D.B. & Coupar, B.E.H. *Nature* **329**, 545-546 (1987).
4. Flexner, C., Hügin, A. & Moss, B. *Nature* **330**, 259-262 (1987).
5. Binnink, J.R., Yewdell, J.W., Smith, G.L. & Moss, B. *J. Virol.* **57**, 786-791 (1986).
6. Pemberton, R.M. *et al. J. gen. Virol.* **68**, 2177-2182 (1987).
7. Pala, P. & Askonas, B.A. *Immunogenetics* **23**, 379-384 (1986).

Brigitte A. Askonas is at the National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK.

Geology

The origins of granites

R. Stephen J. Sparks and Herbert E. Huppert

GRANITIC rocks represent 70 per cent of the exposed surface of the continental crust and many sedimentary and metamorphic rocks are derived from granite by various physical and chemical processes. Between 1785 and 1787 James Hutton of Edinburgh (Fig. 1) made several critical observations on the geology of granite which demonstrated its igneous origin and the intrusive relationship of granites to surrounding rocks. Hutton recognized that these rocks had been melted within the Earth, which was a basis of his seminal book *Theory of the Earth* (1788). This publication initiated modern igneous geology, and how knowledge has progressed since then was the central question at a recent conference* which celebrated the bicentenary of Hutton's book.

It is sobering to reflect that several currently fashionable ideas were anticipated by nineteenth-century geologists, although they often did not have the techniques to validate them (W. Pitcher, Liverpool University). Granite cannot be extracted directly from the mantle by a one-stage process because the mantle generates basalt when it melts. Granite is a product of repeated melting and differentiation in the crust which, over geological time, has concentrated the most fusible and lowest-density components at the Earth's surface (Fig. 2). Some granites form from the melting of metamorphic rocks that have a sedimentary origin, for example, the spectacular leucogranites of the Himalayas that seem to have formed by melting of pelitic schists and gneisses in the hot Tibetan thrust slab during the collision of India and Eurasia 18 million years ago (P. Le Fort, CNRS, Nancy).

Another example is a Peruvian ignimbrite that is essentially an erupted, partially molten pelite with primary igneous muscovite, andalusite, sillimanite and cordierite (M. Pichavant, CNRS, Nancy). The granites of the Lachlan fold belt in south-east Australia reflect the source-rock compositions, either sedimentary (S type) or igneous (I type), that melted and then ascended (B. Chappell, ANU; A. White, La Trobe University). The latter interpretation is supported by ion-probe data (I. Williams, ANU) from inherited zircons with ancient cores that are much older than the granite emplacement, and which record the metamorphic and igneous history of the source rocks.

Different kinds of granites are formed in other situations. Geochemical data from

the Cordilleran batholiths of the western United States indicate there were systematic variations of isotopes, trace elements and mineralogy across these ranges, with rocks becoming increasingly radiogenic inland towards the old Precambrian craton (L. Silver and H. Taylor, Caltech). Attention was drawn to the difficulty of distinguishing young, lower crust as a source of granites from mixtures of mantle-derived magmas and older continental crust (A. Halliday, Michigan University).

Is there an empirical link between the composition of granites and tectonic environment (J. Pearce, University of Newcastle)? Certain trace elements (niobium, rubidium and yttrium) produce very impressive discrimination between granites from syn-collisional environments, arcs, mid-ocean ridges and within-plate environments. Pearce also gave plausible explanations of petrogenetic processes that could be responsible for the different behaviour of these elements.

Although there is agreement on many aspects of granite genesis, there are still some fundamental differences of opinion and perspective. One such issue is the debate on the relative importance of deep-versus shallow-level processes on the petrogenesis of granite and associated volcanic rocks. On the one hand, there are adherents to the US Geological Survey (USGS) standard magma-chamber model who hold that most of the processes dominating the geochemistry and petrological characteristics of silicic igneous rocks occur in large, long-lived, shallow magma chambers. Various geological and geochemical observations on ash-flow tuff caldera systems can be interpreted in

Scottish National Portrait Gallery

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Fig. 1 James Hutton, MD 1726–1797, by Sir Henry Raeburn in about 1780.

terms of this magma-chamber model (P. Lipman, USGS; R. MacDonald, University of Lancaster). At the other end of the spectrum of opinion are those who interpret many rocks in terms of deep-level processes in the source region. The Australian granite school, for example, considers that many granitic magmas involve the mobilization and ascent of partially molten source regions. These regions are emplaced in the upper crust as a mixture of restite crystals and melt, where they freeze as high-level granites or erupt as volcanic rocks. According to this view, the deep-level partial melting processes are much more important. This is not now simply a debate between plutonists and volcanists, as the argument about the petrogenesis of the Fish Canyon Tuff demonstrates (see Stormer, J.C. *et al. J. Petrol.* **28**, 747; 1987).

The physical mechanisms of magma generation and emplacement were addressed in rather few contributions, reflecting the great emphasis on geochemistry in modern granite research, rather than the wider implications of granites for large-scale geodynamic processes. Indeed, there are even some philosophical differences between groups about what is actually meant by 'origin'. The work of the Australian school, for example, has the objective of identifying the age and composition of the deep crustal source regions that melt to generate granites. The identification of the characteristics of the source is one important element of the origin. The cause of the melting, which must have required

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Fig. 2 Basalt dykes cutting the Golmud granite, which was emplaced 260–195 million years ago during the collision of the Qiantang terrane of central Tibet with the Kunlun terrane along the Jinsha suture. (Photograph by C. Lewis.)

* *The Origin of Granites* Heriot-Watt University, Edinburgh, 14–16 September 1987. Organized by the Royal Societies of Edinburgh and London.

large amounts of heat transfer into the crust for the generation of the south-east Australian granites, for example, and the tectonic environment in which this happened, could be argued to be more fundamental to the concept of origin.

What are the emplacement mechanisms of granites? In thermal and fluid-dynamical models of diapir ascent (M. Harrison, State University of New York, Albany), the surrounding rocks are heated to high temperature in a thin skin which then flows around the ascending diapir. Harrison's calculations make it hard to envisage a diapiric rise through the cold upper crust; the diapirs simply ascend too slowly and freeze. Studies of the effect of structural environment on granite emplacement (D. Hutton, Durham University) show that the crust is not just a passive medium through which granites pass, but deviatoric stress conditions and local structural factors can have a major influence on how magmas are emplaced.

When considering the genesis of granite, a fundamental problem is how continental crust is heated sufficiently to generate large amounts of magma. Two main mechanisms have been proposed: first, thickening of crust in collision zones

is followed by conductive heating from below which can be accentuated by deep burial of rocks rich in radioactive elements, a suggestion supported by thermal models in which thickening is accomplished by the stacking of several thrust sheets (E-an Zen, USGS, Virginia); and second, basalt underplating, which brings thermal energy directly into the crust. This second mechanism is attractive as all the main tectonic processes (subduction, extension and plume activity) can involve substantial melting of the mantle and emplacement of basalt beneath or through the crust. We presented a fluid-dynamical model of the emplacement of basalt sills into the crust which predicts that this situation provides a very efficient mechanism for transferring heat between the mantle and crust and for generating substantial volumes of granite. Phenocrysts form in the source region during melting and the resulting granite magma is a mixture of phenocrysts and restite, satisfyingly consistent with the Australian school. □

R. Stephen J. Sparks and Herbert E. Huppert are in the Department of Earth Sciences, University of Cambridge, CB2 3EQ, and the Department of Applied Mathematics and Theoretical Physics, Cambridge CB3 9EW, UK.

Microbial metabolism

Anaerobes pumping iron

Richard B. Frankel

SEVERAL microorganisms, including the bacteria *Leptethrix*, *Siderocapca* and *T. ferrooxidans*¹, use ferrous iron as an electron (energy) source, oxidizing it to ferric iron while respiring O_2 . Some bacteria can use ferric iron at the other end of the electron-transport chain as the terminal electron acceptor. But there have been no indications that the latter is an important process *in vivo*. Now, Lovley and co-workers, on page 252 of this issue², report the isolation and growth of a bacterium from anaerobic mud in the Potomac River basin that oxidizes organic acids to CO_2 while reducing hydrous ferric oxide to magnetite, Fe_3O_4 . This finding has geochemical, evolutionary and palaeomagnetic implications.

The organism described by Lovley *et al.*, cryptically designated GS-15, reduces 8 moles of ferric iron per mole of acetate consumed. The ferric iron in the culture medium is present as an amorphous hydrous iron-oxide precipitate resulting from the hydrolysis of ferric chloride. The production of Fe_3O_4 presumably occurs extracellularly after the export of ferrous ions into the medium, where they subsequently interact with unreduced hydrous ferric oxide. Fe_3O_4 is composed of two ferric and one ferrous iron per formula unit, and apparently cannot be further

reduced by the organism. Hence, only one-third of the ferric iron in the medium is available for respiration. Fe_3O_4 production is nevertheless copious, potentially reaching 1 kilogram per 10 grams of biomass! By comparison, magnetotactic bacteria that use nitrate or oxygen as electron acceptors produce about 0.2 grams Fe_3O_4 per 10 grams of biomass³. Of course, actual Fe_3O_4 production by GS-15 *in vivo* could vary depending on the concentrations of anions such as carbonate, phosphate or sulphide that compete for ferrous ions. Whether the exported ferrous ions are incorporated into Fe_3O_4 or into other iron minerals, it is clear that GS-15 and its relations can have a significant impact on the chemistry of iron in anaerobic sediments. Moreover, because the Fe_3O_4 particles are in the single-magnetic-domain size range, they can have a large effect on the palaeomagnetic intensity of those sediments. Karlin *et al.* recently reported⁴ magnetic evidence for Fe_3O_4 production in suboxic marine sediments, and attributed it to iron reduction by microorganisms. Lovley *et al.* now point out² that ancestors of GS-15 could have played a major role in the formation of Fe_3O_4 in the banded-iron deposits during the Precambrian.

To investigate the chemistry of this

process, Tamaura *et al.*⁵ and Mann⁶ have produced Fe_3O_4 *in vitro* by adding ferrous ions to hydrous ferric-oxide precipitates. The process is thought to involve a solution reprecipitation sequence that begins with the binding of the ferrous ions on the surface of the iron-oxide particles. On the other hand, Lovley *et al.*² did not obtain Fe_3O_4 when they added ferrous ions to their uninoculated medium. This may result from inhibition by acetate or something else, either by chelating the ferrous ions or preventing binding to the surface of the oxide. In viable cultures, acetate would be consumed; that and other changes, such as in pH, could allow the process to proceed. Amorphous hydrous ferric oxide and ferrous ions are known to be precursors to intracellular Fe_3O_4 formation in the bacterium *A. magnetotacticum*⁷.

Lowenstam has distinguished⁸ between biologically induced mineralization (BIM) and matrix-mediated, or boundary-organized, biomineralization (BOB)⁶. In BIM, cellular export of metabolic products leads to extracellular mineral formation with materials in the environment. In BOB, the mineral phases are deposited in preformed organic matrices produced by the organism. Thus, Fe_3O_4 production by GS-15 and *A. magnetotacticum* is biologically induced and matrix-mediated, respectively. In the former, Fe_3O_4 particles have a broad size distribution and do not seem to be associated with an organic matrix, whereas in the latter the particles have a narrow size distribution, definite morphologies and are enveloped by a membrane⁹. Even in the BIM process the dimension of the particles is less than 50 nanometres.

In magnetotactic bacteria, Fe_3O_4 serves as an aid to magnetic orientation and navigation, helping the motile cells to find and remain in the preferred microaerophilic zone¹⁰. For GS-15, Fe_3O_4 could just be a metabolic by-product and have no other biological significance. But these non-motile cells seem to grow in intimate contact with the precipitates in the culture vessel and not in the water column above, which is sensible considering the very low solubility of hydrous ferric oxides and the very high iron requirement of the organism. Fe_3O_4 has a density of 5 and *in vivo* could serve as an anchor for the cells in the habitat where their physiology gives them an advantage over other bacteria. □

1. Ehrlich, H.L. *Geomicrobiology* (Dekker, New York, 1981).
2. Lovley, D.R. *et al.* *Nature* **330**, 252–254 (1987).
3. Blakemore, R.P. *A. Rev. Microbiol.* **36**, 217–238 (1982).
4. Karlin, R. *et al.* *Nature* **326**, 490–493 (1987).
5. Tamaura, Y. *et al.* *J. chem. Soc. Dalton Trans.* 189 (1983).
6. Mann, S. *J. inorg. Biochem.* **28**, 363–371 (1986).
7. Mann, S. *et al.* *Nature* **310**, 405–407 (1984).
8. Lowenstam, H.A. *Science* **244**, 126–130 (1981).
9. Gorby, Y. & Blakemore, R.P. *J. Bact.* (in the press).
10. Spormann, A. & Wolfe, R.S. *FEMS Microbiol. Lett.* **22**, 171–177 (1984).

Richard B. Frankel is at the Francis Bitter National Magnet Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.

Regulation of transcription

An oncogene caught red-handed?

N. J. Short

In the past, DNA-binding oncoproteins have enjoyed a reputation rather like that of Ivan Boesky: they were known to exist and to engage in nefarious activities, but they smoothly eluded all attempts to define exactly what these activities were. In an exciting new development, however, the recently discovered oncoprotein JUN has been subjected to a detailed investigation¹ by Kevin Struhl, leading to its arraignment on the charge (admittedly circumstantial) of impersonating the DNA-binding factor AP-1. Because no version of a factor regulating transcription has ever been shown to be an oncoprotein before, the case for the prosecution is worth examining in some detail.

The oncogene JUN was first found disguised as a retrovirus, loitering under suspicious circumstances in a spontaneous sarcoma of chickens. Many oncogenes have been isolated in this way; vertebrate cells contain numerous genes that are oncogenic if improperly or inaccurately expressed (the so-called cellular or proto-oncogenes), and a retrovirus that has exchanged some of its own DNA for such a gene can often survive if wild-type virus is present to help it to replicate. The virus isolated from the sarcoma in question was indeed shown to be replication-defective and to transform fibroblasts in culture. It was therefore called avian sarcoma virus 17 (ASV17) and tested by dot-blot hybridization for the presence of known oncogenes. No trace of such being discovered, its genome was examined for tracts missing from the wild-type virus, and these were sequenced. The new-found sequences encode a polypeptide of 296 amino acids fused to a viral protein, and the resulting oncogene was given the name of *jun* (from the Japanese for 17)².

So far, the story resembles that of many other better-studied oncogenes. As part of the initiation ceremony for new oncogenes, however, it has for several years been traditional to search protein-sequence databases using a computer, in the hope of finding something resembling the new gene. The results in the present case were quite unexpected; 44 per cent of the amino acids in the carboxy-terminal end of the

JUN oncoprotein are identical to those in the same end of GCN4, a DNA-binding protein from yeast³. GCN4 had already been relatively well studied, and in particular was known to bind DNA via its carboxy-terminal end. Thus, the resemblance between the JUN and GCN4 sequences immediately suggested that JUN is one of the relatively few oncoproteins that are known to exert their effects by binding DNA. Struhl¹ therefore took the bold step of constructing a hybrid gene in which the DNA-binding domain of GCN4 is replaced by its JUN equivalent, placing it on a plasmid in yeast lacking the GCN4 gene, and asking whether the chimaeric gene could fulfil any of the functions of its wild-type counterpart. His audacity paid off; the chimaeric protein, like GCN4, stimulates histidine synthesis when expressed by the yeast (see figure). Admittedly, the hybrid does not work as well as the wild-type protein, and needs an extra domain (apparently to promote dimerization) before it works at all. Nevertheless, as Dr Johnson observed when comparing a female preacher to a dog walking on its hind legs, "The remarkable thing, Sir, is not how well it does it,

but that it can do it at all".

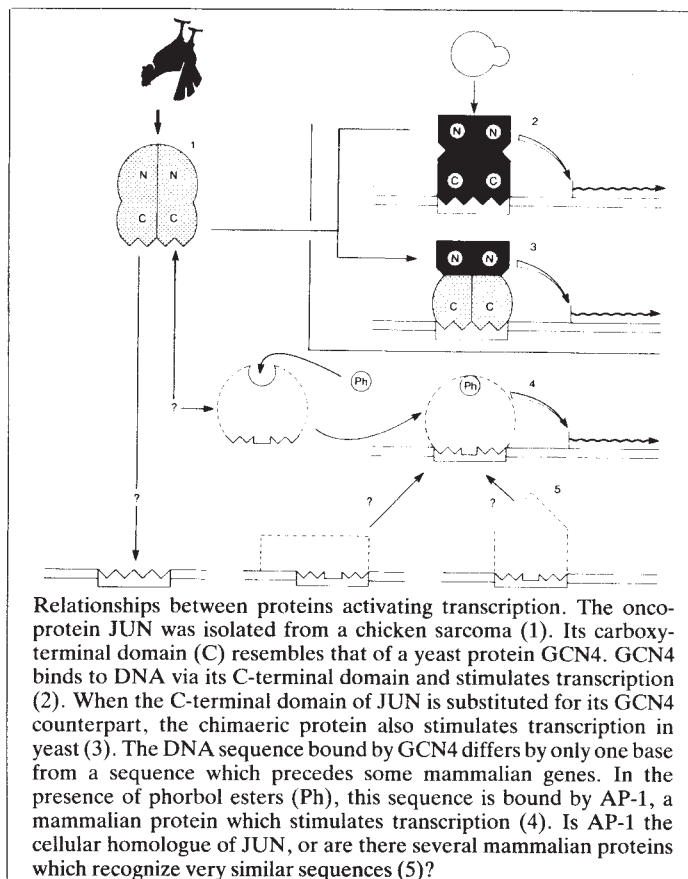
Even more remarkable, however, is the clue that the experiment may give as to the function of the cellular homologue of the JUN proteins. The sequence bound by GCN4 is well known⁴, and bears a striking resemblance to a sequence which appears in front of many mammalian genes whose transcription is induced when the cell is treated with phorbol esters. Such a sequence forms part of the simian virus 40 enhancer and precedes the genes encoding collagenase, stromelysin and metallothionein II_A, among others. This motif is bound by the protein AP-1, which has recently been purified both by Michael Karin's group⁵ and by Robert Tjian's group⁶. Binding by AP-1 stimulates transcription and is increased up to fourfold by phorbol ester treatment. As phorbol esters are notorious tumour promoters, it is highly tempting to conclude that JUN behaves like a permanently activated version of AP-1 and so leads the cell to behave as it if were perpetually exposed to phorbol esters (see figure).

There is no doubt that this hypothesis ties together the known facts about JUN very neatly, and it may even be true. On the other hand, there are several observations suggesting that the real story may not be so simple. The first of these concerns the sequences recognized by GCN4 (or, by extension, JUN) and AP-1. Both motifs are quite well defined, as in each case several genes exist that are known to

bind the proteins in question^{4,5}.

Although the consensus sequences for the two factors differ by only one base in eight, the one-base difference is highly consistent: no site binding AP-1 has a C at position 4, and no site binding GCN4 has a G. Although this may not seem very important, a single base mutation in the recognition sequence is sufficient to prevent the binding of many *trans*-acting factors. A simple extension of the experiments used to define the GCN4 binding site will be needed to show whether the AP-1 motif is acceptable either to GCN4 or to JUN.

A second potential complication concerns the increasing number of motifs involved in the regulation of transcription that have recently been shown to be bound by more than one factor. The paradigmatic sequence showing this behaviour is the so-called octamer motif involved in the control of immunoglobulin and other genes. As long as a year ago this was shown to bind at least two



proteins, whose ratios vary in extracts from different cell lines⁷. The latest publication brings the total to four, of which two may or may not be differentially modified forms of the same polypeptide⁸. Not only do these proteins display subtle differences in their affinities for various versions of their consensus binding site, but some can displace yet other proteins from overlapping sites in lymphoid cells.

This baroque complexity now seems to be the rule rather than the exception; two proteins have recently been described that recognize sequences overlapping the binding site for CTF/NFI⁹ (the protein I described in my previous News and Views article¹⁰), and another recent paper examines four proteins binding to overlapping sites in the simian virus 40 enhancer¹¹. Many of these factors seem to have a tightly controlled tissue specificity, or to be selectively bound and inactivated in some cells, and the multiplicity of proteins therefore probably reflects the highly complex combinatorial system necessary to control gene expression in

multicellular organisms. Nevertheless, it is plainly conceivable that, even if the JUN oncoprotein can bind to the AP-1 motif, other proteins may be able to do so too. It will thus be very interesting to find out whether the cellular homologue of *jun* is the same as the AP-1 that Tjian and others have purified. Even if it is not, the JUN protein will remain interesting for some time to come; as a proven oncoprotein and a near-proven *trans*-acting factor, what genes does it control to produce sarcomas? □

1. Struhl, K. *Cell* **50**, 841–846 (1987).
2. Maki, Y. *et al. Proc natn. Acad. Sci. U.S.A.* **84**, 2848–2852 (1987).
3. Vogt, P.K. *et al. Proc natn. Acad. Sci. U.S.A.* **84**, 3316–3319 (1987).
4. Hill, D.E. *et al. Science* **234**, 451–457 (1986).
5. Angel, P. *et al. Cell* **49**, 729–739 (1987).
6. Lee, W. *et al. Cell* **49**, 741–752 (1987).
7. Staudt, L.M. *et al. Nature* **323**, 640–643 (1986).
8. Rosales, R. *et al. EMBO J.* **6**, 3015–3025 (1987).
9. Dorn, A. *et al. Cell* **50**, 863–872 (1987).
10. Short, N.J. *Nature* **326**, 740–741 (1987).
11. Xiao, J.H. *et al. EMBO J.* **6**, 3005–3013 (1987).

N. J. Short is in the Biophysics Department, King's College London (KQC), 26–29 Drury Lane, London WC2B 5RL, UK.

Cosmology

Wormholes in space-time

Ian Moss

WHEN wormholes appear in the timbers of your home you have something to worry about. When they turn up in empty space then you would not be likely to notice except at the Superconducting Supercollider (SSC) particle accelerator, up for tender in the United States. One of the most important reasons for building the SSC is to find the particle called the Higgs boson that is central to our understanding of the fundamental forces of nature. According to the new results of Stephen Hawking, however, the Higgs particle might never be found if it is an indivisible entity (*Phys. Lett.* **195B**, 337–343; 1987) because of tiny 'wormhole' fluctuations that should exist in the fabric of space as a result of quantum mechanics. Although far too small to be seen directly, these quantum fluctuations prevent Higgs particles from moving anywhere.

The structure of space and time (space-time) includes the relationship of cause and effect, termed the causal structure. Another feature is the interconnectedness of points, called the topology. These

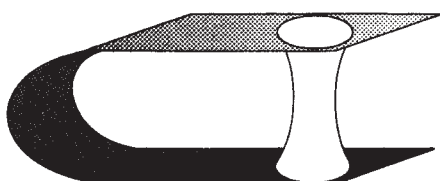


Fig. 1 A wormhole connecting two points on a flat sheet.

and other properties of space-time are subject to the laws of quantum indeterminacy, giving rise to quantum gravity. This is an open and challenging subject, but from what is known about other systems, it is expected that quantum fluctuations are important for distances smaller than the Planck length of 10^{-33} cm.

In Hawking's approach to quantum gravity, the causal structure of space-time is radically different on this scale. The time coordinate becomes replaced by another space coordinate. Fluctuations in the topology of space also occur, connecting points that are otherwise far apart (Fig. 1).

The quantum description of wormholes is similar to that of atoms; for example, both have excited states. A particle can be scattered by a wormhole, leaving the wormhole in its ground state. This process (Fig. 2a) makes a calculable contribution to the mass of the particle. A particle can also be absorbed by a wormhole, forcing the wormhole into an excited state which can then decay to release another particle elsewhere (Fig. 2b). In effect, the particle slips through a wormhole to a distant place. One result of this is that any attempt to produce a Higgs particle will fail, because the wormholes would carry the particle everywhere throughout space, requiring an infinite amount of energy. This leads to a kind of superconductivity of empty space for Higgs particles.

Wormhole states can also interact with

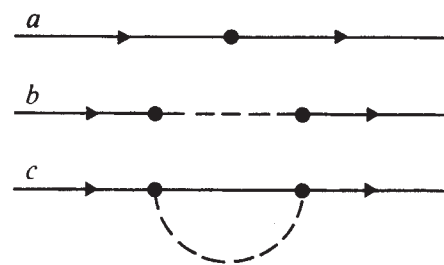


Fig. 2 Interactions of a particle (solid line) with a wormhole (•). a, Collision without excitation of the wormhole; this process contributes to the mass of the particle. b, Collision with excitation (absorption); the particle seems to disappear from one point in space and reappear elsewhere. c, Two wormhole excitations interact to make a 'molecule'.

one another. A simple example is shown in Fig. 2c, where two wormholes form a diatomic molecule joined by a line. These wormholes last only as long as it takes light to travel a Planck length, so that the chemistry of the space-time vacuum is essentially instantaneous.

Besides connecting different points of space, wormholes can also connect parts of the Universe to separate, self-contained pieces of space. This leads to statistical uncertainty in the quantum states of particles beyond that usually associated with quantum mechanics. A similar effect occurs in the theory of black hole evaporation, also devised by Hawking (*Nature* **248**, 30; 1974).

Black holes can be classified as special types of wormholes. Because the states of particles inside a black hole cannot be measured from outside, the information lost defines the black hole's entropy.

Fortunately not all elementary particles are trapped by wormholes. Whether they are depends on their spin, which determines how they look when they are rotated. Particles with zero spin, like the Higgs particle, are affected. Most 'material' particles are fermions, with spin one-half. Wormhole excitations have integer spin, so that they can interact only with even numbers of fermions. Because it is far harder to get two particles near a wormhole than one, the effect is much smaller.

Hawking has constructed an elegant new theory of wormhole excitations, though this is not yet a full quantum theory of wormholes. It seems likely that the best opportunity to test the theory is with the Higgs particle. To do this requires a level of quantization beyond that usually taken in quantum gravity. For historical reasons, this would become third quantization (second quantization arises in Dirac's relativistic quantum mechanics) — but where does the sequence end? The immediate issues can only be decided when the SSC is built and experimental data become available. □

Ian Moss is in the School of Physics, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU, UK.

Cell motility

Movements made visible by microchip technology

Linda Amos

TEN years ago, people were trying to introduce moist specimen chambers into the evacuated columns of electron microscopes as this seemed to be the only way in which cellular organelles an order of magnitude smaller than the wavelength of light could be observed in action. But in the past few years it has been discovered that many 'submicroscopic' motile events can be seen by visible light microscopy, even though the sizes of the organelles cannot be resolved. Initial studies were carried out using fluorescent labels but, more recently, unlabelled structures have been seen with electronic contrast enhancement. As in many other disciplines, recent progress has depended greatly on the response of instrument manufacturers to the demands of leading researchers. The equipment now available for visualizing the movement of cells and their constituents and some of the remarkable results that have been obtained were the subject of a recent meeting*.

Information about many types of cell motility has been gathered using a system devised by the late Robert Allen (see *Nature* 321, 647; 1986). He and his collaborators at Woods Hole Marine Biological Laboratory (MBL) first saw individual microtubules of diameter about 25 nm sliding and squirming over a glass surface, using a high-performance video camera attached to a light microscope. The contrast in high light-level images, such as those obtained by Nomarski differential interference-contrast (DIC) optics, can be greatly enhanced by electronically subtracting a uniform value and amplifying the remaining signal. The Allen system also includes on-line computer-image processing for removing non-uniform background intensity, such as the mottle pattern arising from the lenses, which makes the images even clearer. With this system, it is even possible to see individual microtubule movements within intact cells, as for example in *Allogromia* (J. Travis and S. Bowser, Vassar College), where they appear to slide and flex over the inner surface of the cell membrane (see figure).

Video enhancement has been used chiefly with Nomarski optics, where the image is generated by local changes in mass concentration in the specimen. The totally different image generated by

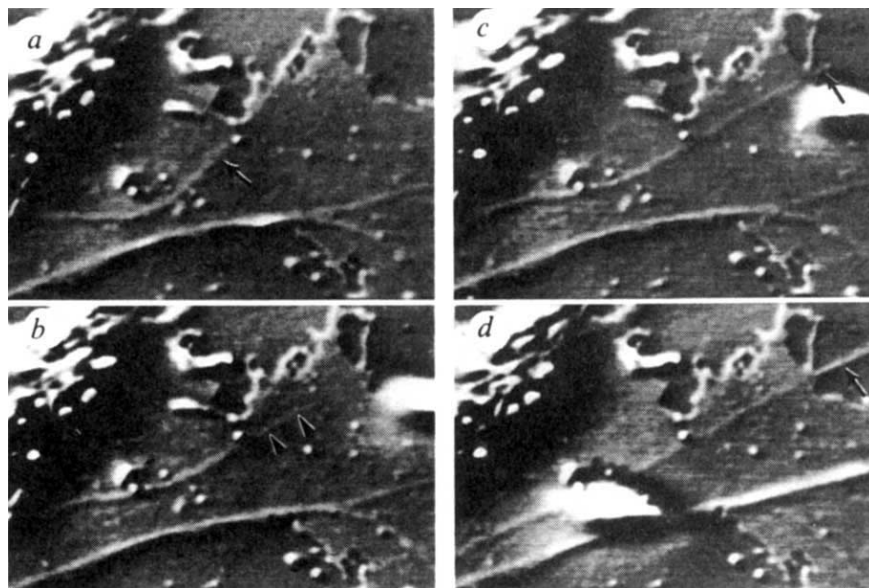
polarizing optics reveals molecular orientation but normally suffers from poor resolution. Combining his innovations in high-resolution polarizing optics with video-enhancement techniques, S. Inoué (MBL) can see details such as single bundles of bacterial flagella and the labile internal structure of fibres from the mitotic spindle. Inoué also demonstrated the advantages of optical disk image storage, not yet available in Britain because of television standards.

Studies of isolated components are providing an explosion of new information, only a smattering of which can be mentioned here. Dynamic instability of microtubules is being studied in detail by observing individual polymers (H. Hotani, Kyoto University; E.D. Salmon, University of North Carolina). And, in the past 2 years, every known ATP-driven molecular motor from cells has been attached to glass and observed moving the appropriate polymer (either actin filaments, fluorescently labelled, or microtubules, imaged directly). Some motors are old favourites, such as myosin subfragment-1 or the flagellar dyneins, others are newly discovered, thanks to this new method of assaying for motility. Perhaps

the ultimate in effective resolution has been achieved in a study of kinesin-coated plastic beads moving along microtubules by B. Schnapp (MBL), who finds a unit step size of only 4 nm, which is the axial spacing of tubulin monomers that comprise the microtubule.

Fluorescence is the most popular way of labelling components that are not otherwise identifiable in the microscope. Video cameras, such as the SIT, ISIT and the even more sensitive VIM, are mainly used to record weak fluorescence signals, although the video-rate signal may have to be integrated over seconds or even minutes to build up a reasonable image. Charge-coupled devices (CCDs) provide a good alternative method of photon counting in such cases. These devices respond linearly to a very large range of intensities, over the complete spectrum of electromagnetic frequencies. It is possible to use as many as five different fluorescent labels on the same specimen and to image the emitted light from each independently (D.L. Taylor, Carnegie-Mellon University). The other essential development for this project is that manufacturers are now designing filters with much greater selectivity than a few years ago. The improved filters even make transmission fluorescence microscopy a feasible alternative to the ubiquitous epifluorescence set-up.

The discussion of confocal laser scanning microscopy suggests there is a very wide range of applications for this technique, which produces optical sections of intact specimens without any background glare. A single spot of laser light, shone



Allogromia filopodia formation by microtubule poking. A microtubule fibril (a, arrow) courses through the lamellipodial cytoplasm and terminates in a short filopod. In b, the fibril has withdrawn from the original filopodium, straightened and begun to extend in a new direction. Note that the newly extended portion of the fibril (arrowheads) appears thinner than the rest of the fibril. c, The tip of the extending fibril has reached the lateral margin of the lamellipodium and poked out a nascent filopod (arrow). At this stage the newly extended portion of the fibril has thickened considerably, so that the entire fibril appears to have a relatively uniform diameter. The new filopod (d, arrow) continues to extend. Note that the fibril has straightened. Time in seconds: a, 0; b, 5; c, 8; d, 27. Magnification, $\times 4,590$. (Courtesy of J.L. Travis and S. Bowser.)

* *Optical Approaches to the Dynamics of Cellular Motility: An MBL Centennial Symposium in Honor of Robert Day Allen* Woods Hole, Massachusetts, 5-8 October 1987.

into the eyepiece of a normal microscope, is made to scan a particular plane of the specimen. A digitized image is built up point by point, by collecting either the fluorescent signal or simply the reflected light that travels back along the same path. Thus, successive levels of fluorescent staining of cell cytoplasm can be imaged independently. Simultaneously scanned (non-confocal) images using

phase contrast or DIC allow the accurate location of the fluorescent components within the cell. As for CCDs, video rates are not practical, but the available rates of 1–5 frames per second are fast enough to compete with any other method of recording fluorescent images. □

Linda Amos is at the MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK.

Radioastronomy

Evading the zone of avoidance

Virginia Trimble

NEARLY one-fifth of the extragalactic sky lurks out of sight, obscured by dust and crowded by images of stars in the disk of our own Milky Way. Astronomers, suspecting with Jane Austen's Catherine, that the locked drawer contains more interesting objects than the open ones, have long wished to penetrate this zone of avoidance with infrared and radio observations. The desire has become more acute with the recognition of very large clumps and holes in the distribution of galaxies and clusters of galaxies^{1,2}. Such large structures are a prime test of galaxy-formation models³. But to trace them out properly, we must be able to map the whole sky. Indeed, at least two known sheet-like features disappear into the zone of avoidance and do not seem to come out the other side⁴.

Technology is, at long last, catching up with wishful thinking, and two publications^{5,6} now report the results of 21-cm radio searches for extragalactic objects at previously inaccessible low galactic latitudes. So far, at least, there be here no dragons. But there are a few dozen newly discovered galaxies — apparently normal spirals, irregulars, and dwarfs, all well outside the Local Group — with the promise of many more to come. No dragons had really been expected, although earlier chance discoveries in the zone of avoidance had included the two large galaxies Maffei I and II immediately outside the Local Group⁷ and, just possibly, a tiny galaxy⁸ almost inside the Milky Way.

The two investigations were rather different in conception. Dow *et al.*⁵ began with a catalogue of infrared sources pinpointed by the Infrared Astronomy Satellite (IRAS). They selected 250 sources in two areas between 3.5° and 15° from the galactic plane, distinguished by particular infrared colours and fluxes to exclude most galactic stars and bits of interstellar cirrus from their list. One area was chosen deliberately to fall between the Perseus–Pisces and Lynx–Ursa Major superclusters of galaxies, where Giovanelli and Haynes⁹ had predicted there should

be a bridge of galaxies with velocities near 5,000 km s⁻¹. Searches for 21-cm emission from neutral hydrogen (H I) in the 250 sources (using the Arecibo 1,000-foot telescope) revealed 63 spiral galaxies with velocities less than 8,000 km s⁻¹, 29 of them previously unknown. And the predicted bridge is there. The average velocity of galaxies in that part of the sky is 4,950 km s⁻¹.

This method has the advantage of considerable efficiency — one knows exactly where to point the radio telescope. It has the disadvantage that IRAS galaxies are largely a somewhat special subset of spirals, rich in gas and young stars. In addition, there still remains a narrower zone of avoidance, because the infrared sources are crowded so close together within 3.5° of the galactic plane that the colours necessary for selection of candidate galaxies cannot be measured. Searches of additional sky area outside this narrower zone can, nevertheless, provide information on many hundreds of currently unknown galaxies and their distribution in space, tracing structure over nearly the whole sky.

Kerr and Henning⁶ adopted a different approach, looking for 21-cm emission without guidance from observations at other wavelengths. In a pilot programme, the 91-m transit radio telescope at the National Radio Astronomy Observatory (NRAO), West Virginia, scanned a random 0.5 per cent of the northern avoidance zone hunting for H I with velocities between 300 and 7,500 km s⁻¹. In more than 100 hours of observing, they found 16 interesting sources. A smaller region of unobserved sky, examined similarly for comparison, yielded 11.

Though the method may not seem very efficient, an H I line profile is remarkably informative once you find it. Spiral galaxies reveal themselves through steep-sided, flat or dimpled-top spectral line shapes, whose central velocity tells you the distance to the galaxy (via Hubble's law). The line width is proportional to the optical luminosity and mass of the galaxy¹⁰. And the intensity of the line

measures the amount of H I present. Emission lines from irregular and dwarf galaxies are more nearly gaussian in shape and rather narrower. Thus Kerr and Henning can say a good deal about their 16 galaxies, even though only one was previously known (as NGC2377) and one other visible on the Palomar Sky Survey prints. Nine have characteristic spiral line profiles and three more may, while the other four are probably dwarf or irregular galaxies. For comparison, the 11 unobscured sources were all previously known and/or identifiable as galaxies and included six normal spirals, four assorted dwarfs and one probable Magellanic irregular. Thus, blind 21-cm searching can pick out a wider range of galaxy types than does the infrared-driven method. In addition, galaxies can be found reliably right down to the galactic equator, leaving no zone of avoidance at all. Once the galaxies have been identified in this way, IRAS data for the relevant locations can be co-added. In the pilot programme, co-adding yielded plausible counterpart sources for most of the obscured galaxies.

Could one probe the entire three steradian zone of avoidance this way? In principle, yes, for there is a suitable southern counterpart to the NRAO 91-m in the Parkes (Australia) 64-m. But the beam width of such telescopes is narrow, about 10°, versus for instance 6° for the 48-inch Schmidt optical telescope that did the Palomar survey. Thus one requires several hundred thousand telescope pointings (of about five minutes each) compared to a few hundred half-hour optical exposures. And only a limited range of velocities can be covered at a time, though one would clearly like to go out to the 12,000 km s⁻¹ limit of existing optical samples.

A complete programme would, therefore, necessarily involve many years and many observers. But, of course, it need not be complete to be useful. In particular, now that both the infrared-driven and blind 21-cm techniques are known to work, deliberate searches can be mounted for continuations and bridges from known clusters and filaments into the zone of avoidance. And we can reasonably expect that at least a few dragon tails will turn up. □

1. Kirshner, R.P., Oemler, A., Schechter, P.L. & Shectman, S. *Astrophys. J.* **248**, L57 (1981).
2. Batusky, D.J. & Burns, J.O. *Astr. J.* **90**, 1413 (1985).
3. White, S.D.M. *et al. Astrophys. J.* **313**, 505 (1987).
4. de Lapparent, V., Geller, M.J. & Huchra, J.P. *Astrophys. J.* **302**, L1 (1986).
5. Dow, M.W., Lu, N.Y., Salpeter, E.E. & Lewis, B.M. *Bull. Am. astr. Soc.* **18**, 1034 (1986).
6. Kerr, F.J. & Henning, P.A. *Astrophys. J.* **320**, L99 (1987).
7. Maffei, P. *Publ. Astr. Soc. Pac.* **80**, 618 (1968).
8. Simonson, S.C. *Astrophys. J.* **201**, L103 (1975).
9. Giovanelli, R. & Haynes, M. *Astr. J.* **90**, 2445 (1985).
10. Tully, R.B. & Fisher, J.R. *Astr. Astrophys.* **54**, 661 (1977).

Virginia Trimble is Professor of Physics at the University of California, Irvine, California 92717 and is Visiting Professor of Astronomy at the University of Maryland, College Park, Maryland 20742, USA.

Of faster brains and bigger teeth

SIR—According to Calvin¹, “slowed human development” is naturally accompanied by slow-growing teeth. It is a common fallacy, however, that the modern human growth period is “slowed”. What characterizes modern humans as unique is a prolongation of the postnatal growth period and not a generalized shift in the rate of growth^{2,3}. Large human brains are energetically expensive to grow⁴ but this calorific expense is offset by prolonged inhibition of somatic growth^{5,7} until brain growth is completed. Early hominids ‘fueled’ enlarging brains either with a reduction in body size, or by prolonging the period of brain growth and delaying somatic growth to later times.

Independent of this, local mechanisms exist whereby teeth of all sizes can grow either quickly or slowly within the constraints of the times available to each tooth during the growth period⁸⁻¹⁰. Beynon and Wood¹¹ show clearly that natural selection favouring large posterior teeth in *Paranthropus* has operated directly upon these local mechanisms. Their fast-growing teeth, then, are not the result of a “long indirect path that confounds straightforward adaptationist reasoning”, as Calvin would have it.

M. CHRISTOPHER DEAN

Department of Anatomy and
Developmental Biology,
University College and Middlesex
School of Medicine,
Gower Street,
London WC1E 6BT, UK

1. Calvin, W.H. *Nature* **328**, 481 (1987).
2. Lovtrup, S. *Syst. Zool.* **27**, 125–130 (1978).
3. Watts, E.S. in *Nonhuman Primate Models for Human Growth and Development* (ed. Watts, E.S.) 41–65 (Liss, New York, 1985).
4. Gould, S. *Ontogeny and Phylogeny* (Harvard University Press, Cambridge, Massachusetts, 1977).
5. Dean, M.C. & Wood, B.A. *Folia Primat.* **43**, 157–180 (1984).
6. Martin, R.D. *52nd James Arthur Lecture* (Am. Mus. Nat. Hist., New York, 1983).
7. Williams, R.F. & Hogden, G.D. *Am. J. Primat. Suppl.* **1**, 181 (1982).
8. Dean, M.C. & Wood, B.A. *Folia Primat.* **36**, 111–127 (1981).
9. Shellis, R.P. *Archs oral Biol.* **29**, 697–705 (1984).
10. Dean M.C. *Am. J. phys. Anthropol.* **68**, 233–238 (1985).
11. Beynon, A.D. & Wood, B.A. *Nature* **326**, 493–496 (1987).

Low levels of aluminium in tea

SIR—Considerable public concern¹⁻³ about the undesirable effects of dietary aluminium has in particular focused on the supposedly high levels of aluminium in tea, reputed⁴ to be 40–100 mg l⁻¹. Recently, we prepared infusions (100 ml water plus 1.0 g tea leaves) of a range of commercially available tea samples from different countries of origin, and measured the aluminium concentration by graphite furnace atomic absorption spectroscopy. The mean concentration of aluminium in

tea made from 12 brands was 3.9 mg l⁻¹, with a range of 2.7–4.9 mg l⁻¹, which is at most only one tenth of the reported levels.

Although our results should allay some of the unease generated by the above publications, there is still no information regarding the bioavailability of aluminium from different dietary sources. Because it is probable that the extent to which aluminium is absorbed from the intestinal tract depends on its chemical form, further research is required to assess the bioavailability of aluminium from tea, compared with other sources in the diet, and to determine the fate of the absorbed aluminium within the body.

SUSAN J. FAIRWEATHER-TAIT
AFRC Institute of Food Research,
Colney Lane,
Norwich NR4 7UA, UK

GEOFFREY R. MOORE

S. E. JEMIL FATEMI

School of Chemical Sciences,
University of East Anglia,
Norwich NR4 7TJ, UK

1. Pearce, F. *New Scient.* **114** (1559), 28 (1987).
2. Kajda, P.K. *B.N.F. Nutr. Bull.* **11**, 149–151 (1986).
3. Hodgkinson, N. *Sunday Times* p.22 (29 February, 1987).
4. Coriat, A.M. & Gillard, R.D. *Nature* **321**, 570 (1986).

Increase of T_c in $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ by exposure to nitrogen

SIR—Recently, Matthews *et al.*¹ have reported an increase in the superconducting transition temperature (T_c) of $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ by as much as 40 K when using nitrogen as a heat exchange medium. In view of the surprising results, and of the fact that this area is riddled with irreproducibilities, we have reinvestigated the effect of nitrogen on the T_c of this ceramic superconductor. Although we also get a nominal increase in T_c by ~ 20 K when measured in a nitrogen atmosphere, we feel that the increase is entirely due to a measurement artefact.

Our samples were made by the usual

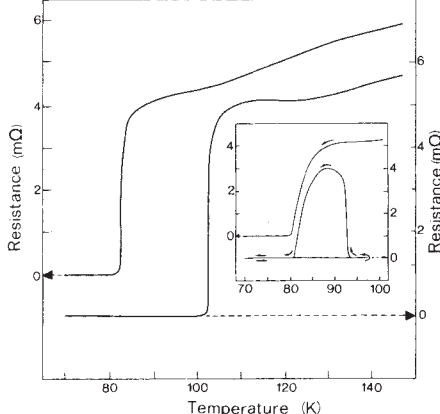


Fig. 1 Resistance versus temperature on cooling ($T_c = 83$ K), and subsequent warming ($T_c = 103$ K). The inset shows identical cooling behaviour, and what happened when we warmed the sample back to 100 K and then re-cooled it (as shown by the arrows).

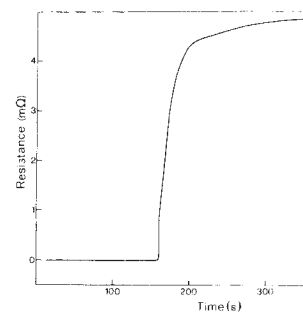


Fig. 2 Resistance versus time, while holding the temperature constant at 93 K. The time at zero corresponds to 300 s after warm-up to 93 K.

ceramic method and the electrical resistivity was measured by the standard four-probe technique. We chose a sample of $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ with $T_c = 83$ K, such that T_c was about midway between the boiling points of N_2 and O_2 . Figure 1 shows the resistance versus temperature for a typical cooling and heating cycle. During cooling, T_c was always found to be 83 K, whereas heating resulted in a T_c of 103 K. Our next measurement consisted of cooling the sample from 140 to 70 K (with $T_c = 83$ K), followed by warming the sample up to 100 K, at which temperature it continued to be in the superconducting state (see Fig. 1 inset). When the sample was cooled again, we found a transition to the normal state at $T_c = 95$ K and, surprisingly, a subsequent re-entry into the superconducting state at 83 K (Fig. 1 inset). This experiment was followed by cooling the sample to 70 K, warming it up to 93 K, where it was still superconducting, and holding the sample at constant temperature (within ± 1 K) for several minutes. The resulting resistance versus time behaviour is shown in Fig. 2. This curve shows that the enhancement of T_c is temporary.

The above results are easily explained by assuming condensation of nitrogen in the sample when it is cooled to 77 K, owing to the intrinsic porosity of the ceramic material used. Depending on the time spent above 77 K, the stored liquid nitrogen keeps the sample below the superconducting transition temperature for some time, even when the environment temperature, monitored by a platinum resistor, is raised considerably above T_c . This leads to the artefact of an apparent increase in T_c . Although this mechanism certainly seems to be operative in our studies, one can only conjecture as to whether this artefact may have been responsible for the observations reported in ref. 1. But we find it significant that the enhancement of T_c in ref. 1 was also observed on the warming cycle and only after condensation of nitrogen. This makes us suspect the same mechanism for the reported “enhancement” of T_c . Before concluding, we note that, in our experiments, the use of oxygen as an exchange

gas did not have any effect on T_c , as the boiling point of O_2 is higher than the T_c of our sample.

D.D. SARMA
C.T. SIMMONS
G. KAINDL

Fachbereich Physik,
Freie Universität Berlin,
D-1000 Berlin 33, FRG

I. Matthews, D.N., Bailey, A., Vaile, R.A., Russell, G.J. & Taylor, K.N.R. *Nature* 328, 786–787 (1987).

MATTHEWS *et al.* reply—The mechanism suggested by Sarma *et al.* was rejected at an early stage in our studies. Simple calculation shows that the heat flow from the outside of a sample maintained at 93 K, say, by exchange gas, to a core of liquid nitrogen at 77 K in the centre of the sample is relatively rapid for the samples used in our studies, which are typically 1–2 cm long and 2×1 mm in cross-section. This heat flow will totally vaporize any stored liquid nitrogen within a few seconds, assuming a thermal conductivity value of $4 \text{ W m}^{-1} \text{ K}^{-1}$ (ref. 1). Consequently, at the heating rates of 0.3 K min^{-1} used by ourselves², equilibrium is maintained throughout the experiment.

The information provided by these authors is inadequate to allow us to estimate the thermal time constant in their case. However, we did carry out an experiment similar to that detailed in their Fig. 2 as part of our original attempts to understand the role played by nitrogen in enhancing T_c . We observed that a stay of 30 min in vacuum, after nitrogen treatment, in the temperature range from ($T_c - 24 \text{ K}$) to T_c , did not affect the enhanced value of T_c (120.5 K) in that experiment. This result is extremely important in that it suggests that the increase in T_c is stable against chamber evacuation provided that $T < 120 \text{ K}$. It also confirms that if any liquid nitrogen condenses within the pores of the sample it is rapidly vaporized and remains in thermal equilibrium with the sample for all temperatures at the heating rate used in this case ($\sim 1 \text{ K min}^{-1}$).

Further evidence for the nature of this enhancement effect comes from a recent careful study of the increase in T_c caused by the controlled addition of nitrogen. As shown in Fig. 1, the continuous exposure of low-density samples to nitrogen leads to a sequence of forms of the $R(T)$ behaviour near to the transition, where R is resistance and T temperature. A model involving the vaporization of stored liquid nitrogen would lead only to changes in T_c and would leave the shape of $R(T)$ at the transition unaltered.

Further evidence, which has been obtained consistently, is the improvement of the critical current/critical field behaviour as nitrogen is absorbed into the specimen¹. Such changes are clearly not consistent with the model proposed.

As the specimen dimensions would

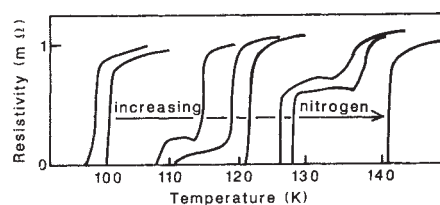


Fig. 1 The resistive transition at various stages in the establishment of a 141-K transition temperature by controlled absorption of nitrogen gas into the sample at low temperatures.

have a direct effect on the vaporization time, and therefore on the observed T_c , if the system were not in equilibrium owing to liquid nitrogen stored in the core, we have made identical observations on two specimens cut from the same sample but with cross-sectional dimensions (r) differing by factors of ~ 3 (corresponding to a factor of nine change in cross-sectional area). Both samples showed identical critical temperatures and $R(r)$ variations both before ($T_c = 90 \text{ K}$) and after nitrogen treatment ($T_c = 111 \text{ K}$). These data also support the view that this effect is not the result of an artefact of the type envisaged by Sarma *et al.*

Despite this weight of evidence, during the past few days we have also carried out a direct measurement of the temperature difference existing in our samples during heating and cooling, using thermocouples

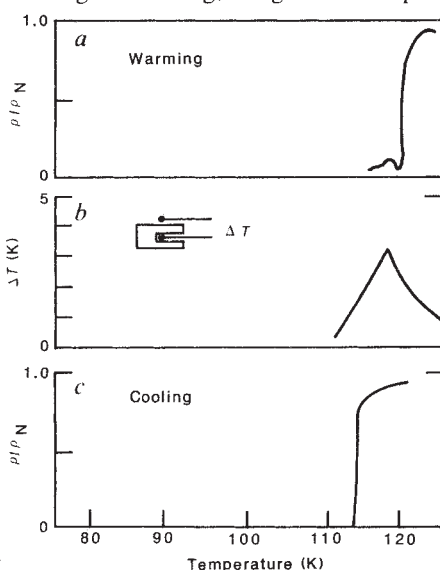


Fig. 2 Changes in internal temperature gradient ($\Delta T \equiv T_{\text{surface}} - T_{\text{core}}$) with temperature during warming and cooling through a nitrogen-enhanced transition. In the warming cycle (a), $T_c = 113 \text{ K}$, and an onset structure appears near T_c ; on cooling (c), T_c is reduced by 5 K, and there is no structure below T_c . ΔT was $< 0.2 \text{ K}$ at all times during cooling. The structure observed during warming at temperatures below the main transition is associated with a Seebeck voltage which appears along the sample, owing to a small temperature gradient parallel to the current direction. The magnitude of this temperature gradient varies with temperature in the same way as the observed transverse ΔT shown in b.

embedded in the superconductor, and correlated these with the electrical properties measured simultaneously. Figure 2 shows the change in R and ΔT ($\equiv T_{\text{surface}} - T_{\text{core}}$) with temperature during a heating and cooling study of the transition. As may be seen (Fig. 2b), a temperature lag (ΔT) of $\sim 3 \text{ K}$ appears close to the resistive onset at T_c during warming. This lag begins at a few degrees below T_c and peaks just above T_c . The major resistive transition (Fig. 2a) occurs just after the peak in ΔT . On cooling (Fig. 2c) there is only a single transition at T_c ; that is, there is hysteresis of $\sim 5 \text{ K}$ in the major transition. In a similar measurement of helium-treated material, for which there was a single transition at $T_c = 118 \text{ K}$, there was again an observable ΔT which also reached its peak value just above T_c . At no time in these experiments did ΔT exceed 4.5 K, and thus no evidence exists for liquid nitrogen in the sample core.

Finally, as it appears that we have indeed observed an intrinsic property of these materials, it is appropriate to try to reinterpret the evidence presented by Sarma *et al.* in their Fig. 2. Contact resistances are typically high in these materials and may lie in the range 1–10 Ω . With a sample current of 10 mA the local heating at the electrodes can be enough to make this region of the sample normal. As the temperature at their measuring point is $0.93 T_c$, the critical current density J_c must be relatively small. The combination of electrode heating and the subsequent temperature rise of the sample will eventually drive the sample normal, as observed by Sarma *et al.*

Unfortunately, in the absence of any details of the experimental technique and contact resistances, it is impossible to do more than suggest that their interesting result is indeed the observation of a nitrogen-enhanced thermal transition at the elevated temperature, driven by the self-heating of the electrodes, combined with effects associated with the critical current density.

K.N.R. TAYLOR
G.J. RUSSELL
D.N. MATTHEWS
A. BAILEY

Advanced Electronic Materials
Technology Group,
School of Physics,
University of New South Wales,
PO Box 1, Kensington, NSW 2033,
Australia

R.A. VAILE
School of Business Technology,
Macarthur Institute of Higher Education,
PO Box 108, Milperra, NSW 2214,
Australia

1. Uher, C. & Kaiser, A.B. *Phys. Rev.* B36(10), (1987).
2. Matthews, D.N., Bailey, A., Vaile, R.A., Russell, G.J. & Taylor, K.N.R. *Nature* 328, 786–787 (1987).
3. Taylor, K.N.R. *et al.* *Phys. Rev.* B (submitted).

Plant breeders' rights and monopoly myths

William Lesser

Legislation may appear to provide a breeding ground for monopoly power and profit among commercial seed propagators, but farmers are canny consumers.

BERLAND and Lewontin¹, in a recent Commentary article, expressed dissatisfaction with Plant Breeders' Rights (PBR) legislation which has been in existence since the early 1960s (1970 in the United States). In essence, their position against this legislation is one of economic efficiency. They see PBR legislation as increasing the monopoly power and profits of breeders while users (farmers) are inhibited from saving seed. Seed saving is, at least in principle, not only less costly for farmers but also economically efficient compared to the annual production, sale and distribution of seed from specialized propagators. Various points support their central argument: the issue of economies of scale in farmer seed processing, the production of new seed varieties that can be promoted as such, the reliance on hybrids, and the emphasis in breeding of traits which are required by PBR legislation as opposed to other ostensibly more productive attributes. My purpose here is to show that many of these concerns — which are widely shared² — are unfounded or largely unrelated to PBR.

Seed saving

It is convenient to begin by examining the system of farmer-saved seeds in the United States and Britain. American farmers save for replanting primarily wheat, cotton and soybean seeds, which on average are replaced every other year³. In Britain about 3 per cent of the soft wheat and barley crop is withheld annually for use as seed. New seed is purchased at those intervals because (1) genetic drift reduces the production potential, and (2) breeding advances make the replacement varieties more productive. For these reasons, it is not conceivable that seed could be efficiently home-produced indefinitely. Nor are there great technical requirements for processing and retaining this seed, except for alfalfa and vegetable seeds, which are very small and difficult to handle.

This pattern is not changed by PBR laws as farmers have exemptions permitting seed saving⁴. However, even if farmers were not permitted to retain seed, the effect on price would not be as great as implied. Suppose farmers bought new seed biennially for \$35 per bushel, supplemented by home-produced seed valued at \$5 per bushel, giving an average annual cost of \$20. Could the seed seller

get \$35 per bushel every year? Possibly not, if the productive value of the seed were not great enough to allow the increased cost. The situation is little different from what a consumer would be willing to pay for a car with an expected life of 100,000 miles compared to one of 50,000 miles.

This argument does not apply to F₁ hybrid seeds, which do not reproduce true-to-type. However, one of the justifications for PBR legislation is to give private breeders incentives to produce self-pollinating varieties. Previously the only form of protection was biological, through hybrids, and research was being directed in that way, an inefficient approach for many crops.

Following the passage of PBR, an increase in private breeding in the United States was documented, most notably for soybeans^{6,7}. Are these expenditures directed to producing seeds little distinguished from previous varieties, the so-called 'cosmetic breeding' argument⁸? This argument is valid only if farmers have no alternative source of information. Such, however, is not the case. Many state and national governments test and rank popular varieties as guides to varietal selection. With such objective information it would be difficult to induce purchases with unsubstantiated sales claims. At the same time, farmers will often try a new variety in a small area before adopting it for use on a broad scale. Recognizing this, in the United States seed companies scatter demonstration plots throughout the main producing areas. With seeds a major productive input, it is not credible that farmers act naively in varietal selection. Indeed Perrin *et al.*⁹ have found a statistical relationship between the PBR and soybean yields in the United States.

Finally, there is no substantiation, at least in the United States, that public research bodies "will have to give up the production of commercial varieties"¹. On the contrary, public breeding of commercial varieties is, and remains, an important function of state agricultural experiment stations⁶. What public plant breeders have done is to protect with PBR their commercial varieties and use the funds to support further research. At Cornell some \$225,000 was raised in 1985 through the commercialization of plant varieties developed in the university¹⁰. Not all varieties are protected, only those with

sufficient commercial potential to justify the expense.

Economics

Several of the economic arguments offered by Berland and Lewontin are puzzling, at best. It is implied that firms purposely raise the cost to thwart entry into breeding. Yet neither evidence nor theory is presented in support. Further, the difficulty of entering breeding should not be overstated, as many firms in the United States found it possible in the era of PBR⁷. Statements about what is hybridized, and why, likewise defy logic and fact. Maize and sorghum were early and leading candidates for hybridization largely because the separate male and female parts make the process relatively inexpensive. Vegetables and flowers must presently be hybridized by hand, while a safe and effective male sterilization method, a necessity for the economical hybridization of wheat, has yet to be discovered. The claim that the yield potential of hybrid maize has not been documented is simply incorrect (see ref. 7).

Berland and Lewontin are correct in thinking that property rights for agricultural plants and seeds do raise public policy issues. Chief among these is assuring the collection, preservation and access to germplasm resources. Regrettably, the points they raise are not among those issues, and their implication that seed breeders are attempting to exploit their newly found property rights to extract monopoly rents from farmers tends to confuse the relevant matters, rather than illuminate them. □

1. Berland, J.P. & Lewontin, R. *Nature* **322**, 785-788 (1986).

2. Claffey, B.A. *Southern J. agr. Econ.* **13**, 29-37 (1981).

3. Leibenluft, R.F. in *Competition in Farm Inputs: An Examination of Four Industries* (US Federal Trade Commission, Office Policy Planning, 1981).

4. In *Yearbook of Agricultural Statistics*, Tables E3, E7, F1, F3 (Eurostat, Statistical Office of the European Communities, Brussels, 1984).

5. Lesser, W. *European Intellectual Property Review* **6**, 172-177 (1987).

6. Butler, L.J. & Marion, B.W. in *The Impacts of Patent Protection in the U.S. Seed Industry and Public Plant Breeding* **16**, 38-39 (University of Wisconsin Monograph NC-117, 1985).

7. Brim, C.A. *Symp. Protection of Biotechnological Inventions*, Ithaca, New York (1987).

8. Fowler, C. & Nelson, G. Hearings before the Subcommittee on Agricultural Research and General Legislation, US Senate (1980).

9. Perrin, R.K., Hunnings, K.A. & Ihnen, L.A. *NCSU. Econ. Res. Rpt* **46**, 36 (1983).

10. Lesser, W. *Int. Review Ind. Property and Copyright Law* **18**, 360-371 (1987).

William Lesser is Associate Professor of Agricultural Economics at Cornell University, Ithaca, New York 14853, USA.

Isotopic constraints on the origin of Hawaiian lavas from the Maui Volcanic Complex, Hawaii

H. B. West*, D. C. Gerlach^{†‡}, W. P. Leeman* & M. O. Garcia[‡]

* Keith–Wiess Geological Laboratories, Rice University, PO Box 1892, Houston, Texas 77251, USA

[†] Department of Earth Sciences, University of Leeds LS2 9JT, UK

[‡] Hawaii Institute of Geophysics, 2525 Correa Road, Honolulu, Hawaii 96822, USA

Strontium, neodymium and lead isotopic data for samples from five volcanoes of the Maui Volcanic Complex, Hawaii, span nearly the total compositional range previously recognized for the Hawaiian islands. These data reveal a systematic compositional continuum for Hawaiian tholeiites which is interpreted to result from partial melting of a variably mixed two-component mantle plume source. In contrast, post-shield-building alkalic cap lavas define distinct compositional trends which appear to result from time- and volume-dependent binary mixing between Hawaiian plume melts and a depleted mantle component.

PETROLOGICAL studies of Hawaiian volcanic rocks have been dominated by research on the island of Hawaii, where there are six major shield volcanoes in various stages of development, including presently active Kilauea and Mauna Loa and the developing Loihi seamount. Because of the relative youthfulness of this island, only three of these volcanoes have reached the alkalic cap stage and none has erupted lavas characteristic of the post-erosional stage¹. Thus, these young volcanoes provide a limited perspective on the chemical evolution of Hawaiian volcanoes. Here we examine spatial and temporal isotopic variations in associated tholeiites, alkalic cap and post-erosional lavas for individual volcanoes from an older part of the Hawaiian volcanic chain.

The present study concerns the Maui Volcanic Complex (MVC), which comprises six major shield volcanoes that are now extinct (West Molokai, East Molokai, Lanai, Kahoolawe and West Maui) or in their final evolutionary stage (Haleakala). These six volcanoes once formed a single island (Fig. 1) that was about half the present area of the island of Hawaii¹. The oldest dated lavas of the MVC are from West Molokai (1.9–2.0 Myr^{2,3}). Ages of shield-building-stage lavas (Table 1) decrease to the southeast³ and the most recent (post-erosional stage) eruption occurred at Haleakala in approximately 1790⁴. The oldest exposed lavas at each of the MVC volcanoes are tholeiites, alkalic cap lavas occur on all but Lanai, and silica-undersaturated post-erosional lavas are found on three (East Molokai, West Maui and Haleakala).

New Sr, Nd and Pb isotopic data are presented here for 18 samples from Kahoolawe, Lanai and West Molokai. Comparable data are available for East and West Molokai, West Maui and Haleakala^{5–10}, and the temporal isotopic evolution of lavas from Haleakala has previously been examined in detail^{9,10}. Brief descriptions of the volcanoes considered in this study are given below; Table 1 gives a summary of ages and rock types exposed at each volcano.

Kahoolawe is subdivided into pre-caldera, caldera-filling and post-caldera lava sequences¹¹. Pre-caldera and caldera-filling lavas are typical Hawaiian tholeiites, similar to those from Kilauea. Alkalic-series lavas are interbedded in the upper caldera-filling section and are found in the post-caldera eruptive sequence¹¹. The thirteen samples analysed span the range from lowermost exposed pre-caldera tholeiites to young post-caldera hawaiites (Table 2).

Lanai lavas are tholeiitic in composition^{11,12}. Field relations, petrography, K–Ar ages and major element analyses of the four samples in this study were reported in ref. 13, trace element analyses in ref. 14.

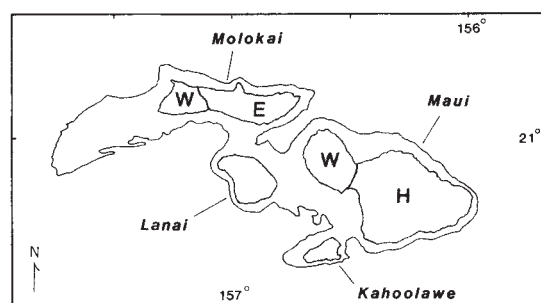


Fig. 1 Map of the Maui Volcanic Complex after ref. 1. Solid line around the islands is the 180-m submarine contour. Molokai comprises two volcanoes: West (W) and East (E). Maui comprises West Maui (W) and Haleakala (H) volcanoes.

West Molokai has received little attention owing to its poorly dissected condition and extensive soil cover; its geology is discussed in ref. 15. In general, the volcano is constructed primarily of typical Hawaiian tholeiites with a thin, discontinuous cap of alkali basalt and hawaiite. We present isotopic data for one tholeiite for which a major element analysis is available¹⁶.

West Maui consists of shield-building (Wailuku Formation), alkalic cap (Honolua Formation) and post-erosional (Lahaina Formation) lavas^{1,17}. Wailuku lavas are predominantly tholeiitic, with interbedded alkali basalts in the uppermost part of the section¹⁸, Honolua lavas range in composition from alkali basalt to trachyte^{16–18}, and Lahaina lavas consist solely of basanitoids¹⁶. Isotopic and major element compositions of lavas from each of these stages have been published^{6–8,16–18}.

Haleakala is deeply dissected and lavas from the shield-building, alkalic cap, and post-erosional stages are well exposed. It is thus ideal for evaluating in detail the petrological and geochemical evolution of a single Hawaiian volcano over a period of at least 0.8 Myr (ref. 3). Previous publications have presented the general geology of Haleakala^{17,19} and compositions of its lavas^{9,10,18,20,21}. Four formations are recognized, including (with decreasing age) Honomanu Formation shield-building tholeiites and minor intercalated alkali basalts, Kumuihihi Formation transitional alkali basalts and hawaiites (exposed only in Haleakala Crater), Kula Formation alkali-series lavas (alkali basalt to trachyte) and Hana Formation post-erosional lavas (basanitoid, alkali basalt and hawaiite).

New Sr, Nd and Pb isotopic results

The lavas of the MVC exhibit substantial ranges in ⁸⁷Sr/⁸⁶Sr (0.70376–0.70442), ²⁰⁶Pb/²⁰⁴Pb (17.712–18.367), ²⁰⁸Pb/²⁰⁴Pb (37.701–38.023) and ¹⁴³Nd/¹⁴⁴Nd (0.51272–0.51298) (Table 2).

[§] Present address: Department of Terrestrial Magnetism, Carnegie Institution of Washington, 5241 Broad Branch Road, Washington, DC 20015, USA.

In fact, the ranges in $^{87}\text{Sr}/^{86}\text{Sr}$ and $^{143}\text{Nd}/^{144}\text{Nd}$ ratios (Fig. 2) are twice those for tholeiites from the island of Hawaii and span virtually the entire compositional spectrum observed for the Hawaiian islands. The overlap of Lanai and Kahoolawe data with those for Koolau (Oahu) tholeiites clearly demonstrates that the latter volcano is not "anomalous" with respect to other Hawaiian volcanoes²². Despite the wide compositional range observed for Kahoolawe lavas, there is no obvious systematic temporal variation, and pre-caldera lavas alone encompass the entire compositional range. This observation contrasts with the systematic temporal Sr-Nd-Pb isotopic variations characteristic of both shield-building and alkalic cap lavas from Haleakala^{9,10}. It is worth noting that at Haleakala, isotopic compositions of post-erosional lavas fall entirely within the field for alkalic cap lavas.

Figure 3 shows the lead isotope systematics of MVC and other Hawaiian volcanoes. Kahoolawe lavas display a wide range in $^{206}\text{Pb}/^{204}\text{Pb}$, but relatively narrow ranges in $^{207}\text{Pb}/^{204}\text{Pb}$ and $^{208}\text{Pb}/^{204}\text{Pb}$. Kahoolawe lead isotopic compositions overlap those of Koolau lavas, and Lanai tholeiites contain the least radiogenic lead yet reported for Hawaiian lavas. The lead isotopic composition of the West Molokai tholeiite is similar to those determined for post-erosional Honolulu Series (Oahu) lavas, but it has distinctly higher $^{87}\text{Sr}/^{86}\text{Sr}$ and lower $^{143}\text{Nd}/^{144}\text{Nd}$ ratios.

Hawaiian magma source mixing systematics

Combined Sr-Pb systematics for Hawaiian lavas define an approximately triangular field with apices represented by tholeiites from Koolau-Lanai-Kahoolawe and Kilauea, respectively, and by Koolau post-erosional lavas from Kauai (Fig. 3). These data are consistent with mixing between a minimum of three compositionally distinct end-members, each represented potentially by the compositional extremes of the Hawaiian field. Within the Hawaiian field, two distinct trends are recognized (Fig. 3): a shield-building (primarily tholeiite) trend, characterized by a negative Sr-Pb correlation, that is best exemplified by Kahoolawe, and a post-shield-building (alkalic cap/post-erosional) trend, characterized by a positive Sr-Pb correlation, represented by Haleakala (shaded region in Fig. 3 and (to a lesser extent) West Maui).

The Sr-Pb array for shield-building lavas ('SB' in Fig. 4) is consistent with binary mixing between an end-member with

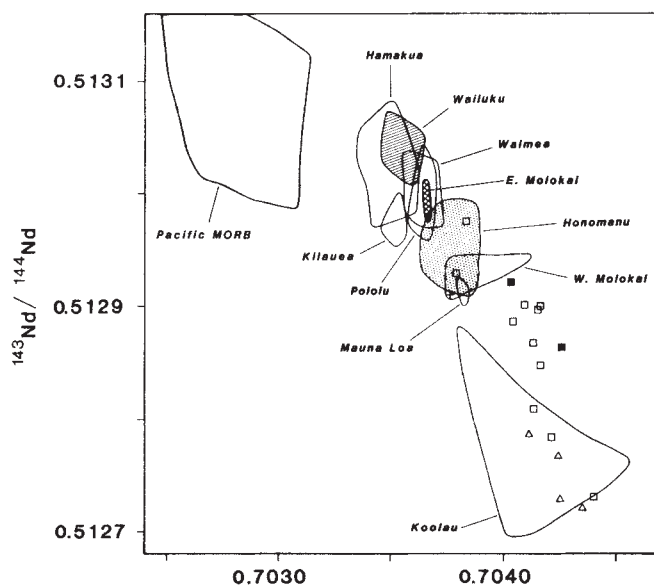


Fig. 2 $^{143}\text{Nd}/^{144}\text{Nd}$ versus $^{87}\text{Sr}/^{86}\text{Sr}$ for Hawaiian tholeiites, including Kahoolawe (tholeiites, open squares; hawaiites, solid squares), Lanai (open triangles) and West Molokai (cross). Fields for other MVC tholeiites are shaded as follows: Wailuku (West Maui), diagonal stripes; East Molokai, cross-hatched; Honomanu (Haleakala), heavy stippling; West Molokai, light shading. A field representing Pacific MORB (mid-ocean-ridge basalt) is shown for comparison. Additional Hawaiian and MORB data are from the literature (see ref. 7 for listing of Hawaiian isotope data references). Other Hawaiian shield-building formations (and their respective volcanoes) are as follows: Hamakua (Mauna Kea), Pololu (Kohala) and Waimea (Kauai). Fields for Waimea and Hamakua are based on unpublished data from D.C.G. and W.P.L. and F. A. Frey *et al.* (in preparation), respectively.

relatively radiogenic strontium and unradiogenic lead and another with relatively unradiogenic strontium and radiogenic lead. The post-shield-building array is also consistent with binary mixing, but differs in that one end-member must lie along the shield-building array whereas the other must have relatively unradiogenic strontium and lead. These mixing arrays are distinct and have opposite slopes, and therefore provide strong evidence that shield-building and post-shield-building lavas cannot have been generated from a common heterogeneous source by different degrees of mixing between source components. Instead, tholeiites and alkalic cap lavas appear to have been generated from separate, two-component sources. In addition, the strong correlation of data within these two arrays is consistent with binary mixing between end-members of relatively restricted isotopic composition, although curvatures of the arrays suggest that these end-members probably have different Sr/Pb ratios.

The three isotopically distinct end-members suggested by the Hawaiian Sr-Pb data can be construed as the following: (1) a component with a 'primitive mantle' signature (PM, essentially equivalent to 'bulk silicate Earth'); (2) 'enriched mantle' (EM); and (3) 'depleted mantle' (DM, essentially a MORB (mid-ocean-ridge basalt)-signature component, possibly representing depleted mantle). These end-members are similar to the Koolau, Kea and PE end-members of ref. 7, and correspond in broad terms, to the PUM (or BSE), HIMU and DMM components of ref. 23. If end-member compositions remained relatively constant beneath the MVC during the past 2.3 Myr then isotopic compositions of MVC lavas must reflect variations in mixing proportions between these end-members. Figure 4 shows a mixing grid constructed to examine the mixing systematics of MVC lavas in the context of these hypothetical end-members.

Data for Kahoolawe, which span half the shield-building

Table 1 Ages and rock types represented at individual MVC volcanoes

Volcano/formation	Age (Myr)	Stage	Rock types*
West Molokai			
Lower member	1.52-1.99 ^{2,3}	Shield-building	Thol
Upper member	?	Alkalic cap	AB, Haw
East Molokai			
Lower member	1.51-2.00 ^{2,3}	Shield-building	Thol, AB
Upper member	1.31-1.46 ²	Alkalic cap	Haw to Tr
Kalaupapa	0.34-1.24 ^{3,13}	Post-erosional	AB, Bas
Lanai			
Lanai volcanic series	1.20-1.46 ⁵	Shield-building	Thol
Kahoolawe			
Pre-caldera	?	Shield-building	Thol
Caldera-filling	?	Shield-building	Thol, AB
Post-caldera	1.02-1.05 ³	Alkalic cap	AB, Haw
West Maui			
Wailuku	1.31-1.97 ^{2,3}	Shield-building	Thol, AB
Honolua	1.20-1.50 ^{2,3}	Alkalic cap	AB to Tr
Lahaina	1.30 ³	Post-erosional	Bas
Haleakala (East Maui)			
Honomanu	0.83 ³	Shield-building	Thol, AB
Kumulihi	0.70-0.91 ³	Transitional	AB, Haw
Kula	0.36-0.88 ^{2,3}	Alkalic cap	AB to Tr, Bas
Hana	0-?*	Post-erosional	Bas, AB, Haw

* Rock types abbreviated as follows: tholeiite (Thol), alkali basalt (AB), hawaiite (Haw), trachyte (Tr), basanitoid (Bas).

Table 2 Isotopic compositions of lavas from the Maui Volcanic Complex

Sample	Rock type*	⁸⁷ Sr/ ⁸⁶ Sr†	¹⁴³ Nd/ ¹⁴⁴ Nd‡	²⁰⁶ Pb/ ²⁰⁴ Pb§	²⁰⁷ Pb/ ²⁰⁴ Pb	²⁰⁸ Pb/ ²⁰⁴ Pb
Kahoolawe						
Pre-caldera						
KW-24	Thol	0.704131 (8)	0.512868 (20)	18.025	15.429	37.759
KW-25	Thol	0.704162 (8)	0.512848 (14)	18.047	15.428	37.786
KW-1	Thol	0.704136 (8)	0.512809 (20)	17.954	15.445	37.805
KW-2	Thol	0.704211 (10)	0.512784 (20)	17.921	15.439	37.733
KW-7	Thol	0.703833 (10)	0.512975 (10)	18.337	15.450	37.990
KW-6	Thol	0.704039 (11)¶	0.512887 (18)¶	18.125¶	15.433¶	37.849¶
KW-5	Thol	0.704223 (10)	0.512902 (20)	18.120	15.431	37.842
		0.703785 (12)¶	0.512929 (20)¶	18.369¶	15.458¶	38.030¶
		0.704609 (8)	0.512928 (20)	18.364	15.451	38.016
Caldera-filling						
KW-23	Thol	0.704090 (8)	0.512901 (16)	18.036	15.430	37.800
KW-19	Thol	0.704399 (12)¶	0.512731 (15)¶	17.939¶	15.453¶	37.842¶
KW-16	Thol	0.704419 (12)	0.512741 (14)	17.946	15.454	37.836
		0.704149 (12)	0.512897 (16)	18.149	15.445	37.845
		0.704159 (10)	0.512900 (12)	18.092	15.439	37.866
Post-caldera						
H1440	Haw	0.704257 (10)	0.512864 (18)	18.005	15.447	37.817
KW-14	Haw	0.704032 (10)	0.512924 (16)	18.027	15.466	37.770
			0.512917 (13)			
Lanai						
OX067	Thol	0.704249 (8)	0.512729 (20)	17.854	15.435	37.727
				17.853	15.420	37.702
OX068	Thol	0.704111 (8)	0.512784 (17)	17.871	15.436	37.701
OX069	Thol	0.704352 (8)	0.512721 (18)	17.886	15.431	37.742
OX078	Thol	0.704239 (10)¶	0.512768 (20)	17.712	15.428	37.738
				17.713	15.424	37.745
West Molokai (lower member)						
C-162	Thol	0.703758 (10)	0.512910 (18)	18.071	15.444	37.731

* Abbreviations as in Table 1.

† Normalized to $^{86}\text{Sr}/^{88}\text{Sr} = 0.1194$; E&A standard = 0.70800.‡ Normalized to $^{146}\text{Nd}/^{144}\text{Nd} = 0.7219$; BCR-1 standard = 0.512640.

§ Normalized and corrected for mass discrimination based on results for NBS981. Uncertainty ~ 0.05% per a.m.u..

¶ Analysis of leached sample (2.5 N HCl decanted several times until colourless).

Sr-Pb array, reveal that tholeiites from individual Hawaiian volcanoes are isotopically diverse and therefore cannot represent homogenized melts (as suggested in ref. 24). The shield-building array could result from mixing within a two-component (PM-EM) 'mantle plume' source, which on melting produces tholeiitic magmas. Mixing between isotopically distinct magmas cannot be ruled out, but this would not alleviate the requirement of at least two distinct tholeiite magma sources beneath Hawaii. The absence of systematic temporal isotopic variations in Kahoolawe tholeiites rules out progressive introduction or exhaustion of one component and implies that the proportions of PM and EM in the plume melts are variable even over relatively short time spans. This effect could result from inherent source variations in PM:EM ratio, differences in solidus temperatures for EM versus PM components (as would result from enriched blobs²⁵ or streaks²⁶ in the Hawaiian plume source) or possibly from short-term variations in the volume or composition of an introduced enriched metasomatic component (see, for example, ref. 27). There is no correlation between the age of tholeiites from a particular MVC volcano and position along the Hawaiian Sr-Pb (Fig. 3) or Sr-Nd (Fig. 2) array; therefore, Hawaiian shield-building lavas as a whole cannot have been generated by progressive mixing between PM and EM components over a longer timescale (such as the age of the Hawaiian islands).

Minimal involvement of DM-type material in the genesis of Hawaiian tholeiites is inferred from the absence of trends toward that component (Fig. 4). This inference is consistent with high melt production rates during the shield-building stage, with consequent rapid magma ascent and eruption, and with limited

residence times in depleted upper mantle and MORB crust, all of which would tend to minimize contamination by MORB-composition wall-rock.

The transition for MVC volcanoes (specifically, West Maui and Haleakala) from shield-building to alkalic cap activity appears to be accompanied by a shift to lower $^{87}\text{Sr}/^{86}\text{Sr}$ and higher $^{143}\text{Nd}/^{144}\text{Nd}$ and $^{206}\text{Pb}/^{204}\text{Pb}^{8-10}$. The cause for this shift is not clear; if it reflects an increase in the EM:PM ratio of the post-shield-building source, then the shift could result either from a decrease in melt production coupled with relatively enhanced proportions of more readily fusible EM streaks or blobs, or from an increase in involvement of an enriched metasomatic component. Alternatively, alkalic cap lavas may be derived from separate, subsidiary plumes that were not involved previously in the generation of shield-building magmas¹⁰. The latter scenario obviates the need to explain production of incompatible trace element-enriched alkalic cap lavas from a source that should be depleted in those elements following the extraction of voluminous shield-building magmas.

The post-shield-building array ('PSB' in Fig. 4) is consistent with mixing between Hawaiian plume magmas, with compositions that lie along the shield-building array, and a DM component. Also, because the post-shield-building array has temporal significance, a progressive incorporation of the DM component is inferred^{9,10}. Increasing involvement of this component could reflect progressive contamination of either the plume source or plume-derived magmas by depleted uppermost asthenosphere. For example, decreased melt production in the plume during the waning stages of magmatic activity could promote increased melt stagnation and wall-rock interaction

(for example, at the lithosphere–asthenosphere boundary). In addition, relatively small-volume subsidiary plumes might be more susceptible to contamination than would larger-volume plumes associated with primary shield-building activity.

Haleakala post-shield-building lavas also display an apparent systematic temporal decrease in EM:PM ratio (Fig. 4)¹⁰. This trend could result from progressive exhaustion of available EM material in the post-shield-building magma source if a finite mantle volume is involved in melt production, and/or may reflect the influence of small compositional heterogeneities in the DM reservoir.

Implications for Hawaiian mantle components

Sr, Nd and Pb isotope systematics of ocean island basalts (OIB) imply that the sub-oceanic mantle sources of these lavas have retained isotopic heterogeneities for times on the order of 10^9 yr^{29,30}. But it is not known whether the isotopic variability observed in OIB reflects the existence of a broad spectrum of end-member compositions or, alternatively, varied mixing proportions between relatively few distinct end-members. It is unlikely that diverse, time-dependent processes such as convective source mixing, extraction or intrusion of silicate melts, or infiltration of non-silicate metasomatic fluids would uniformly affect large mantle domains, and it is reasonable to expect the existence of isotopic heterogeneities on a variety of scales. Although there is considerable leeway regarding the compositions and spatial distribution of the mantle reservoirs inferred for Hawaii, our interpretations above are not dependent on the exact choice of end-member compositions.

There has been considerable debate over the possible existence of primitive mantle (see, for example, refs 23, 31–34). In this paper, the PM component is taken to represent a mantle reservoir possessing the inferred isotopic composition of primitive mantle (see also ref. 25). It has been proposed^{32–34} that this PM component cannot actually be primitive (chondritic) mantle because oceanic basalts (MORB and OIB), including those from Koolau which have PM-like isotopic compositions, possess non-chondritic and constant Nb/U, Nb/Th and Pb/Ce ratios. But elemental ratios measured in lavas may not be representative of their source mantle, particularly if a multi-component source is involved or complex melting³⁵ and fractionation³⁶ processes modify these ratios. For example, negative correlations between Nb/Th and elemental abundances (including rare-earth elements) in Kahoolawe tholeiites imply that this ratio was controlled by magmatic processes. Furthermore, Kahoolawe Nb/Th ratios span a considerably greater range (9.6 – 23.5) than that proposed for Hawaiian volcanoes (15.2 ± 1.7)³⁴, and approach chondritic values (7.9 – 8.8). Although trace element ratios cannot at present rule out the involvement of primitive mantle, the debate over the origin of the PM component in Hawaiian lavas is by no means resolved.

Similarly, there is uncertainty concerning the origin of a Hawaiian EM component. It could represent previously depleted (possibly MORB-source) mantle that was subsequently re-enriched by recycling of oceanic crust and lithosphere^{32–34} or sediment³⁷, magma injection, or metasomatic processes resulting in increased U/Pb, and to a lesser extent Rb/Sr and Nd/Sm. In fact, the range in isotopic compositions for Pacific MORB may reflect addition of an EM-like component to a DM reservoir (see ref. 38).

The EM composition assumed here is constrained only to the extent that it cannot have less radiogenic lead than that in the most radiogenic Hawaiian lavas. Lavas from certain oceanic islands (such as St Helena, Tubuai and Mangaia^{39,40}) are characterized by even more radiogenic $^{206}\text{Pb}/^{204}\text{Pb}$ (>21.6)⁴⁰, and have strontium and neodymium isotopic compositions between those of our PM and DM end-members. Presumably these lavas were generated, at least in part, from mantle domains that underwent more extreme U/Pb enrichment than our postulated EM end-member. If such extreme EM compositions are more

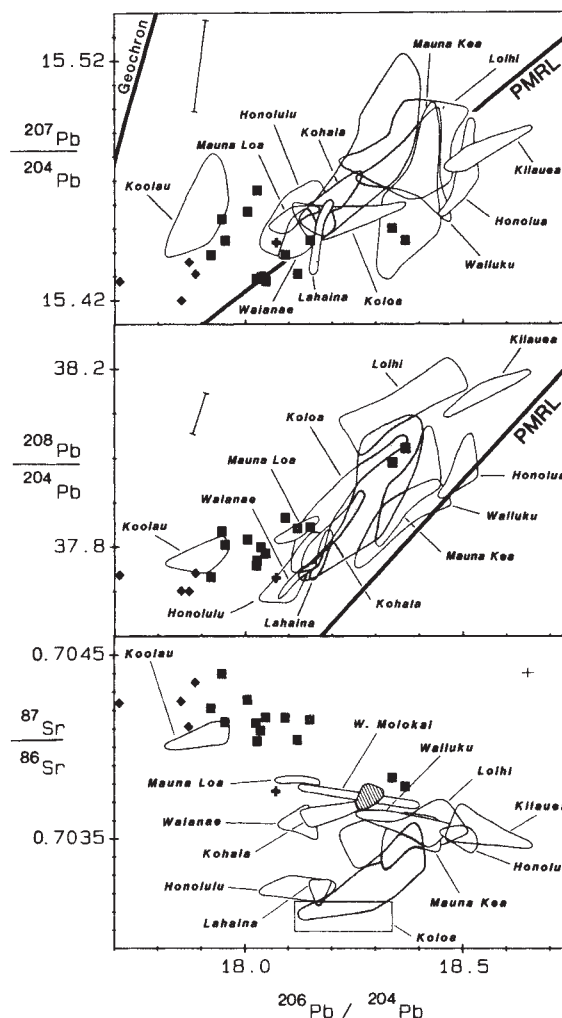


Fig. 3 Pb–Pb and Sr–Pb isotope plots for lavas from Kahoolawe (squares), Lanai (diamonds) and West Molokai (cross). Shaded area denotes the field for Haleakala post-shield-building lavas (Kumuiahi, Kula and Hana Formations), which become progressively less radiogenic with age¹⁰. Diagonally striped field in Sr–Pb figure represents field for Haleakala shield-building lavas (Honomanu Formation)^{9,10}. Indicated formations and their respective volcanoes as in Fig. 2, with the following additions: Honolulu (alkalic cap) and Lahaina (post-erosional), West Maui; Koolau (post-erosional), Kauai; Honolulu (post-erosional), Koolau. Data sources as in Fig. 2. Geochron and Pacific MORB regression line (PMRL, ref. 10) are shown for reference.

appropriate for Hawaii than the example used here, then Kilauea tholeiites could contain a commensurately smaller proportion of EM component than shown in Fig. 4. Comparable enrichments in rubidium and neodymium would be less effective in accelerating isotopic evolution than would additions of uranium, owing to the buffering effect of large strontium and neodymium abundances in the mantle relative to that of lead³⁷, and to more rapid decay of ^{235}U and ^{238}U compared to ^{87}Rb and ^{147}Sm . If these enrichments occurred sufficiently long ago, it would have been possible to generate a component with relatively radiogenic lead but with strontium and neodymium compositions between those of DM and PM. But because the age, degree and nature of enrichment are likely to vary throughout the mantle, a compositional continuum of EM reservoirs probably exists on a global scale.

It is unlikely that EM represents an enriched portion of heterogeneous MORB lithosphere, nor can Hawaiian isotope systematics represent simple two-component mixing of PM-type material with such lithosphere²⁸. In order to reproduce the

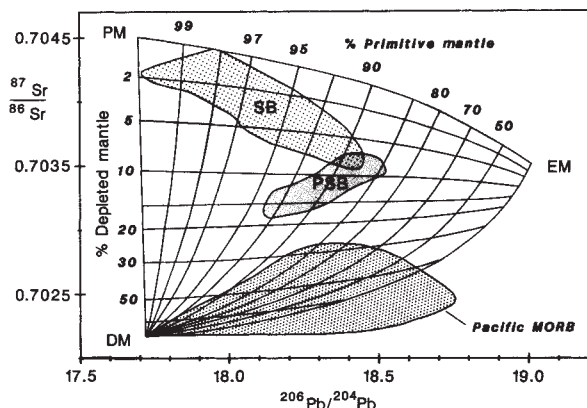


Fig. 4 Sr-Pb isotopic mixing model for the Maui Volcanic Complex (MVC). SB and PSB fields denote MVC shield-building and post-shield-building lavas, respectively. Data sources as in Fig. 2. End-member compositions from ref. 10. EM, enriched mantle; PM, primitive mantle; DM, depleted mantle. Upper scale refers to the percentage of PM in a PM-EM mix (the Hawaiian plume composition); vertical scale refers to the percentage of DM in a DM-plume mix. See text for details.

oppositely correlated Sr-Pb shield-building and post-shield-building arrays, such a model requires that shield-building magmas interact primarily with enriched-type lithosphere whereas post-shield-building magmas interact primarily with strongly depleted lithosphere. There is, however, no compelling explanation for why tholeiites and alkalic cap magmas should interact exclusively with the compositional extremes of this lithosphere (avoiding lithosphere of intermediate Sr-Pb isotopic composition), nor why one magma suite should interact with only one compositional end-member and not the other. Mixing between a relatively fixed end-member (PM) and a heterogeneous end-member (MORB lithosphere) would probably not produce the observed binary mixing trends; rather, such interaction would produce a fan-shaped field (as in ref. 28), which is not observed (see Fig. 4).

For several reasons, we consider it improbable that contamination of magmas at shallow depths by altered oceanic crust was significant in producing the shield-building Sr-Pb array (Fig. 3). First, although uptake of uranium from sea water⁴¹ could result in old, altered oceanic crust having radiogenic lead compositions similar to our postulated EM component, such material would also possess excessively radiogenic strontium⁴². Second, the low neodymium content of sea water⁴¹ makes it

unlikely that seawater alteration would produce significant neodymium heterogeneities in oceanic crust. Thus, strong correlations between neodymium and the other isotopic systems in Hawaiian basalts suggest that neither their strontium nor their lead compositions have been affected significantly by interaction of these magmas with altered oceanic crust. In addition, the relatively restricted EM composition inferred is inconsistent with isotopically heterogeneous⁴³ oceanic crust. Finally, the seismicity associated with recent eruptions at Kilauea and Mauna Loa^{44,45} indicates that Hawaiian tholeiites ascend rapidly from upper mantle depths (>60 km⁴⁶), with only short residence times (<100 yr⁴⁶) at crustal depths. Under these conditions, crustal contamination appears to have been negligible, as evidenced by insignificant strontium and neodymium isotopic variations during the 1969-1973 Mauna Ulu (Kilauea) eruptions⁴⁷.

Scale of mantle heterogeneity

A possible inverse relationship between magma production rates and isotopic diversity in oceanic basalts (MORB and OIB) has led to the proposal⁴⁸⁻⁵⁰ that basalt compositions are controlled by variable degrees of melting of mantle sources that are heterogeneous on a small scale. In this model, MORB is generated by large and OIB by small degrees of melting of a common heterogeneous mantle; homogeneous melts would be produced in cases where the scale of melting is larger than the scale of heterogeneity. If small-scale heterogeneities are responsible for the isotopic diversity of Hawaiian lavas, and if tholeiites are generated by larger degrees of melting than are associated alkalic lavas, then it follows that tholeiites should be comparatively homogeneous. Because they clearly are not, we infer that the scale of mantle heterogeneity in the shield-building plume source (PM+EM) is larger than the scale of melting. Thus, the data for MVC shield-building lavas imply large-scale heterogeneities in the Hawaiian plume source, and the isotopic heterogeneity of post-shield-building lavas requires a distinct petrogenesis, involving interaction between shield-building-type magmas and a DM (MORB-signature) component (see ref. 10).

We thank W. M. White and A. W. Hofmann for their thoughtful and constructive reviews. D.C.G. acknowledges a postdoctoral fellowship from the University of Leeds and thanks R. Cliff, P. Guise and R. Green for technical support. H.B.W. and W.P.L. thank D. & W. Kennedy for invaluable contributions to our personal QIE. We thank the U.S. Navy for providing access and transportation to the island of Kahoolawe and M. Manghnani for providing the West Molokai sample from the Macdonald collection. This research was partly supported by NSF grants to W.P.L. and M.O.G.

Received 27 February; accepted 23 September 1987.

- Macdonald, G. A., Abbott, A. T. & Peterson, F. L. in *Volcanoes in the Sea* (University of Hawaii Press, Honolulu, 1983).
- McDougall, I. *Geol. Soc. Am. Bull.* **75**, 107-128 (1964).
- Naughton, J. J., Macdonald, G. A. & Greenberg, V. A. *J. Volcan. geotherm. Res.* **7**, 339-355 (1980).
- Oostdam, B. L. *Geol. Soc. Am. Bull.* **76**, 393-394 (1965).
- Clague, D. A. *et al. Pacif. Sci.* **36**, 411-420 (1982).
- Stille, P., Unruh, D. M. & Tatsumoto, M. *Nature* **304**, 25-29 (1983).
- Stille, P., Unruh, D. M. & Tatsumoto, M. *Geochim. cosmochim. Acta* **50**, 2303-2319 (1986).
- Hegner, E., Unruh, D. M. & Tatsumoto, M. *Nature* **319**, 478-480 (1986).
- Chen, C.-Y. & Frey, F. A. *J. geophys. Res.* **90**, 8743-8768 (1985).
- West, H. B. & Leeman, W. P. *Earth planet. Sci. Lett.* **84**, 211-225 (1987).
- Stearns, H. T. *Hawaii Div. Hydrogr. Bull.* No. 6 (1940).
- Wentworth, C. K. B. *Bishop Mus. Bull.* No. 24 (1925).
- Bonhommet, N., Beeson, M. H. & Dalrymple, G. B. *Geol. Soc. Am. Bull.* **88**, 1282-1286 (1977).
- Budahn, J. R. & Schmitt, R. A. *Geochim. cosmochim. Acta* **49**, 67-87 (1985).
- Stearns, H. T. & Macdonald, G. A. *Hawaii Div. Hydrogr. Bull.* No. 11 (1947).
- Macdonald, G. A. *Mem. Geol. Soc. Am.* **116**, 477-522 (1968).
- Stearns, H. T. & Macdonald, G. A. *Hawaii Div. Hydrogr. Bull.* No. 7 (1942).
- Macdonald, G. A. & Katsura, T. *J. Petrology* **5**, 82-133 (1964).
- Macdonald, G. A. *U.S. geol. Surv. misc. Invest. Ser. Map* I-1088 (1978).
- Macdonald, G. A. & Powers, H. A. *Geol. Soc. Am. Bull.* **57**, 115-124 (1946).
- Macdonald, G. A. & Powers, H. A. *Geol. Soc. Am. Bull.* **79**, 877-888 (1968).
- Roden, M. F., Frey, F. A. & Clague, D. A. *Earth planet. Sci. Lett.* **69**, 141-158 (1984).
- Zindler, A. & Hart, S. R. A. *Rev. Earth planet. Sci.* **14**, 493-571 (1986).
- Staudigel, H. *et al. Earth planet. Sci. Lett.* **69**, 13-29 (1984).
- Allegre, C. J., Hamelin, B. & Dupré, B. *Earth planet. Sci. Lett.* **71**, 71-84 (1984).

- Olsen, P., Yuen, D. A. & Balsiger, D. *Phys. Earth planet. Int.* **36**, 291-304 (1984).
- Menzies, M. A. & Murthy, V. R. *Am. J. Sci.* **280-A**, 622-638 (1980).
- Chen, C.-Y. *Nature* **327**, 49-52 (1987).
- Sun, S.-S. & Hanson, G. N. *Geology* **3**, 297-302 (1975).
- Tatsumoto, M. *Earth planet. Sci. Lett.* **38**, 62-87 (1978).
- White, W. M. *Geology* **13**, 115-118 (1985).
- Hofmann, A. W., Jochum, K. P., Seifert, M. & White, W. M. *Earth planet. Sci. Lett.* **79**, 33-45 (1986).
- Newsom, H. E., White, W. M., Jochum, K. P. & Hofmann, A. W. *Earth planet. Sci. Lett.* **80**, 299-313 (1986).
- Hofmann, A. W. *Chem. Geol.* **57**, 17-30 (1986).
- O'Hara, M. J. *Nature* **314**, 58-62 (1985).
- O'Hara, M. J. *Nature* **266**, 503-507 (1977).
- Armstrong, R. L. *Phil. Trans. R. Soc. A* **301**, 443-472 (1981).
- LeRoex, A. P. in *Mantle Metasomatism* (eds Menzies, M. A. & Hawkesworth, C. J.) 389-422 (Academic, London, 1987).
- Sun, S.-S. *Phil. Trans. R. Soc. A* **297**, 409-445 (1980).
- Palacz, A. Z. & Saunders, A. D. *Earth planet. Sci. Lett.* **79**, 270-280 (1986).
- Michard, A., Albarède, F., Michard, G., Minster, J. F. & Charlou, J. L. *Nature* **303**, 795-797 (1983).
- Menzies, M. A. & Seyfried, W. E. *Earth planet. Sci. Lett.* **44**, 463-472 (1979).
- Hart, S. R. & Staudigel, H. *Earth planet. Sci. Lett.* **58**, 202-212 (1982).
- Ryan, M. P., Koyanagi, R. Y. & Fiske, R. S. *J. geophys. Res.* **86**, 7111-7129 (1981).
- Dzurisen, D., Koyanagi, R. Y. & English, T. T. *J. Volcan. geotherm. Res.* **21**, 177-206 (1984).
- Wright, T. L. *J. geophys. Res.* **89**, 3233-3252 (1984).
- Hofmann, A. W., Feigenson, M. D. & Raczek, I. *Contr. Miner. Petrol.* **88**, 24-35 (1984).
- Batiza, R. *Geophys. Res. Lett.* **8**, 853-856 (1981).
- Batiza, R. *Nature* **309**, 440-441 (1984).
- Macdougall, J. D. & Lugmair, G. W. *Nature* **313**, 209-211 (1985).

The structure of an oligo(dA)·oligo(dT) tract and its biological implications

Hillary C. M. Nelson, John T. Finch, Bonaventura F. Luisi & Aaron Klug

Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

Poly(dA)·poly(dT) has unusual properties in that it cannot associate into nucleosomes and short, phased runs of it cause DNA bending. The crystal structure of a B-type DNA dodecamer containing a homopolymeric run of six A·T base pairs shows that this region possesses special structural features, including a system of bifurcated hydrogen bonds, which explains some of the properties of this simple homopolymer.

THE homopolymer poly(dA)·poly(dT) has many unusual and distinct structural features that differentiate it from other B-type DNA sequences. The helical repeat in solution was determined to be 10.0 base pairs (bp) per turn, in contrast to 10.5 bp per turn found for random DNA and alternating poly(dA)·poly(dT)^{1,2}, and the axial rise of the helix is less than that found for other B-type DNA fibres (3.2 versus 3.4 Å)³. Fibres of the homopolymer are not affected by the environmental changes such as humidity, cation, or salt concentration, which induce fibres of other sequences to undergo the B- to A-type transition⁴. Neither poly(dA)·poly(dT) nor poly(dG)·poly(dC) can be reassociated into nucleosomes, but the latter case is explained by the rigidity conferred by the three hydrogen bonds per base pair^{5,6}. What is it about poly(dA)·poly(dT) that produces the same apparent rigidity, as well as the resistance to environmental changes?

Because of its distinctive properties, poly(dA)·poly(dT) is important in dictating both the translational and rotational positioning of DNA on nucleosomes. Lengths of 80–100 bp are sufficient to be at least partially excluded from nucleosomes^{7,8}, whereas runs of about 20 bp are preferentially positioned at the ends of the nucleosome core⁸. An instance when the translational positioning of nucleosomes may be of functional significance is found in yeast: oligo(dA)·oligo(dT) tracts act as upstream promoter elements that constitutively activate transcription, perhaps due to the exclusion of nucleosomes from that region of the chromosome⁹. Shorter lengths can be incorporated into nucleosomal DNA, but even then, runs more than 4 or 5 bp long have preferential rotational settings on the nucleosome supercoil at positions of minimum curvature^{10,11}.

A further example of the use of the poly(dA)·poly(dT) structure in biological recognition and control is the ability of short, repeated runs of the polymer to influence the configuration of long runs of free DNA. Runs of about 5 bp of the homopolymer phased every 10 to 11 bp (every helical repeat) cause an overall bend to the DNA configuration, as seen by abnormally slow gel electrophoretic mobility^{12,14}, relaxation times in electric dichroism^{15,16}, and electron microscopy¹⁷. The relationship between the altered DNA configuration and the repeating sequence pattern of oligo(dA)·oligo(dT) was first observed in trypanosome kinetoplast minicircles^{18,19}, and has now been shown to occur in eukaryotic and prokaryotic replication origins and in transcriptional regulatory regions^{20–23}. The bent configuration is required for activity at some of these sites: for example, SV40 T-antigen and DNA interaction is dependent on the altered DNA configuration²⁴.

The questions we have addressed are: what is the structure of poly(dA)·poly(dT), and can this structure explain the unusual physical properties and hence its biological function? We have solved by x-ray analysis the crystal structure of a DNA dodecamer whose sequence includes six bps of oligo(dA)·oligo(dT). We show that the oligo(dA)·oligo(dT)

C1	G2	C3	A4	A5	A6	A7	A8	A9	G10	C11	G12
G24	C23	G22	T21	T20	T19	T18	T17	T16	C15	G14	C13

Fig. 1 Sequence used for crystallization. The base pairs are C1–G24, G2–C23, and so on and the base steps defined as C1, G2–C23, G24. The single strands (dCGCAAAAAGCG) and (dCGCTTTTTCGCG) were synthesised on an Applied Biosystems DNA synthesizer, and purified individually by column and gel chromatography as an ammonium salt. The strands were then mixed together at equimolar concentrations; more than 99% of the DNA was double-stranded as estimated by autoradiography of a gel-electrophoresed sample. The needle-shaped crystals (average size 0.16 × 0.22 × 0.4 mm) were grown from solutions containing 0.2 mM DNA, 10.0 mM magnesium acetate, 0.5 mM spermine-HCl, and 5% (v/v) 2-methyl-2,4-pentanediol (MPD), vapour diffused at 4 °C against 30 to 45% MPD. X-ray data were collected on a Watts–Hilger four-circle diffractometer for three crystals at 4 °C. Over 5,000 reflections were collected in the resolution range of 20–2.5 Å, and the intensities corrected for Lorentz polarization, absorption and time-dependent decay. Symmetry-equivalent reflections were averaged; and the data from the three crystals were merged with an *R*-merge in intensity of 10%. Of the final 2,572 unique reflections in the 20–2.5 Å resolution range, 2,093 reflections had amplitudes greater than 0.5σ; data below 0.5σ were excluded from the refinement. In the 3–2.5 Å resolution range, two thirds of the reflections were used in the final data set.

tract is essentially straight, the base pairs being parallel to each other and perpendicular to the helix axis. The base pairs do have an unusual structure, due to the high propeller twist (see below) at each A·T base pair. This results in maximal overlap of the bases on each strand and the formation of a run of additional, non-Watson–Crick, cross-strand hydrogen bonds. These distinctive structural features explain what differentiates poly(dA)·poly(dT) from other DNA sequences, why poly(dA)·poly(dT) seems conformationally rigid, and suggest how runs of oligo(dA)·oligo(dT) are involved in DNA bending.

Structure determination and refinement

The sequence used is shown in Fig. 1, and the crystallization methods are described in the legend. The crystals diffract to 2.5 Å, but only weakly beyond 3 Å, and have unit cell dimensions of *a* = 25.4, *b* = 40.7, *c* = 65.8 Å, with space group P2₁2₁2₁. Our approach was to use a sequence related to the self-complementary CGCGAATTCGCG dodecamer, whose structure had already been solved^{25,26}. This was justified by the similarities in the patterns of X-ray intensities, indicating the same crystal packing. It made possible a solution of the structure by using the sugar–phosphate chains of the CGCGAATTCGCG dodecamer as a starting model and manually substituting our sequence for that of the other dodecamer. However, it also meant that much care was needed to efface the original structure during the refinement (details in legend to Fig. 2).

Fig. 2 Stereo view of 'partial' difference map for base pair A6-T7 in the final refined structure. This map is of the kind used in the refinement. They are generated by omitting one nucleotide pair from the calculation of structure factors, and using these phases to calculate a difference $|F_{\text{obs}}| - |F_{\text{calc}}|$ map. The density at the position of the omitted nucleotide pair is then entirely due to its contribution to the X-ray amplitudes. The map shown is contoured at $0.17 \text{ e}/\text{\AA}^3$, twice the value of the r.m.s. density. Because our sequence

is not self-complementary, the structure is polar and can be introduced into the unit cell either way up. The two possible models, differing in polarity, were refined against the X-ray data, using the rigid body refinement program CORELS (ref. 41). At first, the entire model was defined as one rigid body, varying the resolution in successively higher resolution shells until reaching the 10–3 Å range. At this point, the individual nucleotides were treated as rigid bodies, and then after *R*-factor convergence, the constituent sugars, bases and phosphate groups were allowed positional freedom. The resolution shell was extended to the 8–2.5 Å range, and the refinement continued. The $2|F_{\text{obs}}| - |F_{\text{calc}}|$ and $|F_{\text{obs}}| - |F_{\text{calc}}|$ electron density maps were examined for satisfactory convergence at each stage of the refinement, using an Evans and Sutherland graphics machine with the program FRODO (ref. 42) (as modified by P. Evans). When the *R*-factor fell from 40% to about 30%, the two models were compared. One model consistently had a *R*-factor that was lower by 1–2% than the model in the opposite orientation. A comparison of these partial difference maps, generated as described above, showed that with the better choice of orientation, the purine-rich and pyrimidine-rich strands could be discriminated, whereas this was not so in the opposite orientation. The refinement was continued with only the better model. In order to minimize bias in the phases from the starting model, extensive use was made of partial maps and recalculation of phases. Partial difference maps were generated omitting each base pair in turn and these were used as a guide, along with difference maps calculated with all atoms present, to adjust by computer graphics any change in the position of the base pairs. After two rounds of this procedure, with the phases regenerated after each step, the model was refined using the Hendrickson-Konnert refinement program NUCLSQ, in which positions and isotropic temperature factors of individual atoms are refined by least squares, subject to a set of distance and general geometry restraints^{43,44}. No restraints were applied to hydrogen bond distances for the central eight base pairs or to the sugar-phosphate torsion angles. Four such cycles of generating partial maps, fitting the base pairs into the electron density, and refining the fitted structure were carried out. After the first cycle, solvent molecules were located where stereochemically feasible if the peaks were present in the $|F_{\text{obs}}| - |F_{\text{calc}}|$ and $2|F_{\text{obs}}| - |F_{\text{calc}}|$ maps before and after density fitting and phase recalculations. All peaks were assigned to water molecules on the basis of their refined temperature factors and proximity to polar groups; however, at this resolution, we cannot distinguish between oxygen and other light ions. A total of 27 water molecules were included in the final model. There were no unassigned peaks in the $|F_{\text{obs}}| - |F_{\text{calc}}|$ map that were greater than $0.23 \text{ e}/\text{\AA}^3$, which is four times the value of the r.m.s. density.

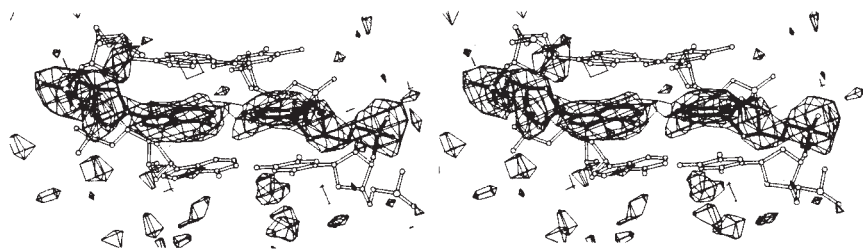


Fig. 3 Stereo view of the structure, the A-T base pairs being shown with black bonds, and the G-C base pairs with white bonds. The best helix axis for the oligo(dA)·oligo(dT) tract was determined using the program HELIX (refs 27, 29), and is shown in the figure. The helical repeat averages 10.0 bp per turn. Although the helical structure of this region is similar to B-type DNA, there are several differences from the standard fibre structure which are a consequence of its high propeller twisted structure. These differences include the decreased axial rise ($3.1 \pm 0.1 \text{ \AA}$ compared with $3.4 \text{ \AA}$ for fibre B-DNA) and a decreased minor groove width, as narrow as $9.1 \text{ \AA}$ at the centre of the AT region, compared with $12 \text{ \AA}$ for fibre B-DNA. Note that the major groove width is only slightly narrower than the average B-DNA values ($17.2 \text{ \AA}$ at the narrowest versus an average of $17.5 \text{ \AA}$ for fibre B-DNA). As has been found in other DNA crystal studies, the sugar-phosphate torsion angles can vary greatly, suggesting an inherent flexibility that will accommodate structural changes^{27–30}. The distribution of angles of the structure presented here fall within the range observed for other DNA crystals. It is interesting to note that all the average torsion angles on the thymine strand are somewhat different and vary more widely than those on the adenine strand. Although the AT region shows some heteronormity, the differences are not as extreme as that modelled for the heteronomous fibre form of poly(dA)·poly(dT) (ref. 45). The average δ torsion angles, defined as $\text{C5}'\text{-C4}'\text{-C3}'\text{-O3}'$, are $128^\circ (\pm 16)$ and $116^\circ (\pm 29)$ respectively for the adenine and thymine strands; this suggests sugar puckers around C1'-exo and C2'-endo, similar to what has been found for other B-type crystal structures. The χ torsion angles, defined as $\text{O4}'\text{-C1}'\text{-N1-C2}$ for pyrimidines and as $\text{O4}'\text{-C1}'\text{-N9-C4}$ for purines, are $-101^\circ (\pm 8)$ and $-105^\circ (\pm 20)$ respectively for the adenine and thymine strands, showing that the bases have similar orientations around their glycosidic bonds.

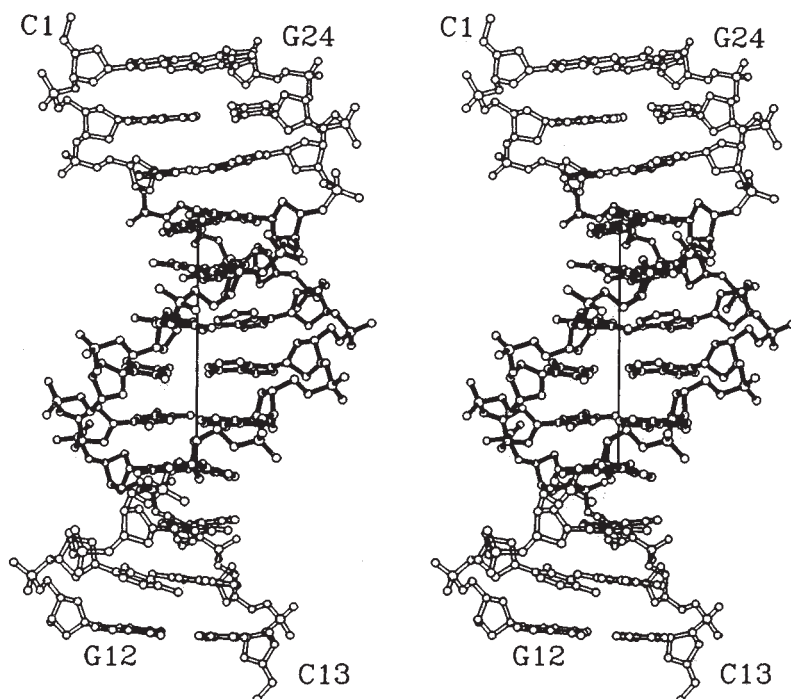


Table 1 Helical parameters of a DNA dodecamer with a run of six A-T base pairs

Base step	Roll (deg.)	Tilt (deg.)	Slide (Å)	Twist (deg.)	Rise (Å)	Pr tw (deg.)	Buckle (deg.)
C1-G24	-1.1	-2.2	-0.2	42.2	2.9	19.0	2.4
G2-C23	-2.8	3.0	-0.4	38.1	3.0	12.3	-1.1
C3-G22	10.3	1.7	-0.9	29.4	4.2	7.7	-1.1
A4-T21	1.0	-1.2	0.1	35.6	3.2	14.6	0.4
A5-T20	-0.8	-0.7	0.1	35.4	3.2	23.0	-2.0
A6-T19	1.5	0.8	0.7	38.5	2.9	25.5	-0.8
A7-T18	-0.1	0.2	-0.2	33.2	3.2	22.5	5.5
A8-T17	-5.1	-1.9	0.2	38.0	3.0	17.7	-2.9
A9-T16	3.9	0.2	-0.9	34.6	3.9	19.2	5.0
G10-C15	-10.0	-4.1	-0.7	36.6	3.1	10.8	-3.7
C11-G14	8.8	2.6	-0.6	42.0	3.6	15.2	-5.2
G12-C13						6.2	-0.1

Pr tw, propeller twist. The helical parameters are calculated using the programs BROLL and CYLIN (refs 27, 29). Because the two bases of a base pair are often non-coplanar, the mean plane of the base pair is used when appropriate for the calculations. Roll, tilt and slide are defined here for one base pair relative to a successive base pair, so they refer to the geometry of the base step. Roll is the rotation of the long axis of a base pair in the direction of the major or minor groove (by convention positive when rotating away from the minor groove). Tilt is the rotation around the short axis towards either sugar phosphate backbone (by convention here positive if rotating towards the adenine strand). Slide is the relative displacement of the C6-C8 long axis of one base pair with respect to the successive one (by convention positive when towards the major groove). Twist and rise are defined with respect to the helix axis, although they are relatively insensitive to the choice of a best-helix axis. Twist is the rotation between successive base pairs about the best-helix axis. Rise is the average axial distance between C1' atoms on successive base pairs. Propeller twist and buckle are defined for each base pair as described in the text. By convention, propeller twist is positive for the counter-clockwise rotation of the distant base when viewed from the near base along the sugar-sugar vector. By convention here, buckle is positive when opening towards the 3' end of the adenine strand.

The final model, shown in Fig. 3, has an *R* factor of 20% for the 2,003 reflections greater than 0.5σ in intensity in the 8 to 2.5 Å resolution range. The r.m.s. deviation in atomic positions between the starting model and the final model is 2.3 Å, and indicates that the starting model had been effectively effaced. The thermal parameters are within observed ranges for other DNA crystal structures, implying that there is no extra disorder within the crystals^{27,28}. Indeed, the average temperature factors for the adenines and thymine were actually lower by several Å² than for the guanines and cytosines. The individual isotropic temperature factors range from 13 to 57 Å², with averages over atoms within groups of 39 for the phosphate, 33 for the sugars, and 25 for the bases. The solvents had individual isotropic temperature factors ranging from 28 to 71, with an average of 49 Å².

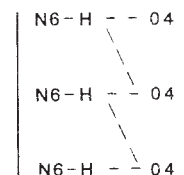
The crystal packing is similar to that found for the native dodecamer, brominated derivatives, and mismatch variants, in which the minor grooves of symmetry-related molecules in the unit cell overlap by three base pairs at their termini²⁷⁻²⁹. There are strong contacts between the bases of G2 and G12 in symmetry-related molecules, and weaker contacts between the bases of G14 and G24. The other contacts involve sugar-phosphate atoms, and presumably have a minor effect in determining base position, given the inherent flexibility of the sugar-phosphate chain. Thus, the last two base pairs are the most affected by crystal packing forces, whereas the oligo(dA)·oligo(dT) region is relatively unaffected by crystal packing forces.

Propeller twist

One of the most obvious features of the oligo(dA)·oligo(dT) region of the structure is that the bases within a base pair are non-coplanar, as seen in Figs 3 and 4. This feature was recog-

nized in the electron density maps in the early stages of refinement; therefore specific Watson-Crick hydrogen bond restraints were not employed in this region in refining the structure. Despite their omission from the refinement, they emerged in the final model with average distances well within the standard values of 2.7 to 3.0 Å (see legend to Fig. 4). The non-coplanarity has two principal components: 'propeller twist' and 'buckle'^{27,29}. Propeller twist is defined as the rotation of the bases along their longitudinal axis, and buckle is the dihedral angle of the bases along their short axis after the propeller twist has been flattened out by rotating to zero. As seen in Table 1, the buckle is variable, ranging from -2.9 to 5.5°, but averages to an almost planar 0.9°. The propeller twist is large, ranging from 15 to 25° with an average of 20°, even higher than the 17° propeller twist seen for the AT base pairs in the CGCGAATTCGCG structure³⁰.

The high degree of propeller twist has two effects on the structure of oligo(dA)·oligo(dT), both of which improve the overall stability of the helix: it helps maximize purine-purine stacking interactions and creates a potential system of additional hydrogen bonds. The propeller twist causes the major groove side of each base to point towards the 3' end of its strand. As seen in Fig. 4, this pushes the N6 atom of adenine, which is in the major groove, towards the O4 atom of thymine on the 3' side of it, making feasible a non-Watson-Crick hydrogen bond diagonally across the major groove, in addition to the normal Watson-Crick hydrogen bonds. The same proton from the amino group would thus be close to two hydrogen bond acceptor atoms (carbonyl oxygens), therefore allowing the formation of a three-centred bifurcated hydrogen bond^{31,32}. As carbonyl oxygens can receive two hydrogen bonds as well, the overall effect on the oligo(dA)·oligo(dT) region is to produce a zig-zag system of hydrogen bonds down the major groove.



It would also seem to follow that a minimum of three A-T base pairs is needed to stabilize the interaction, by allowing the middle base pair to make two of these three-centred bifurcated hydrogen bonds. The cross-strand diagonal hydrogen bonds are long, but the geometry of the interaction, with the combination of long and short bonds, is typical for these types of bifurcated hydrogen bonds, which have been located in proteins³³, as well as in tRNA (ref. 34). In DNA, from geometrical and chemical restraints, a system of cross-strand diagonal hydrogen bonds in the major groove can form only in certain homopolymer stretches of B-type DNA, where the hydrogen bonds of the base pair straddle the helix axis. Isolated instances of cross-strand diagonal hydrogen bonds have been found in the minor groove at the G·A mismatch positions of a decamer structure³⁵, and can be found in the major groove at the AA steps of some of the brominated derivatives of the CGCGAATTCGCG dodecamer.

Base-stacking, specifically purine-purine stacking, is a very significant force in stabilizing homopolymer structures, and poly(dA)·poly(dT) maximizes this through the high degree of propeller twist which rotates the bases around their longitudinal axis, allowing more interactions between neighbouring bases. As seen in Fig. 5, the adenines stack so as to overlap their six-membered rings and the thymine stack such that the C5-methyl group makes weak carbon-carbon van der Waals' contacts with C6 atoms of the pyrimidine ring below. This, together with the additional hydrogen bonding system, causes an increase in the helical twist and a decrease in the helical rise (see Table 1). We cannot estimate the relative contributions of the base stacking and additional hydrogen bonds to the overall stability

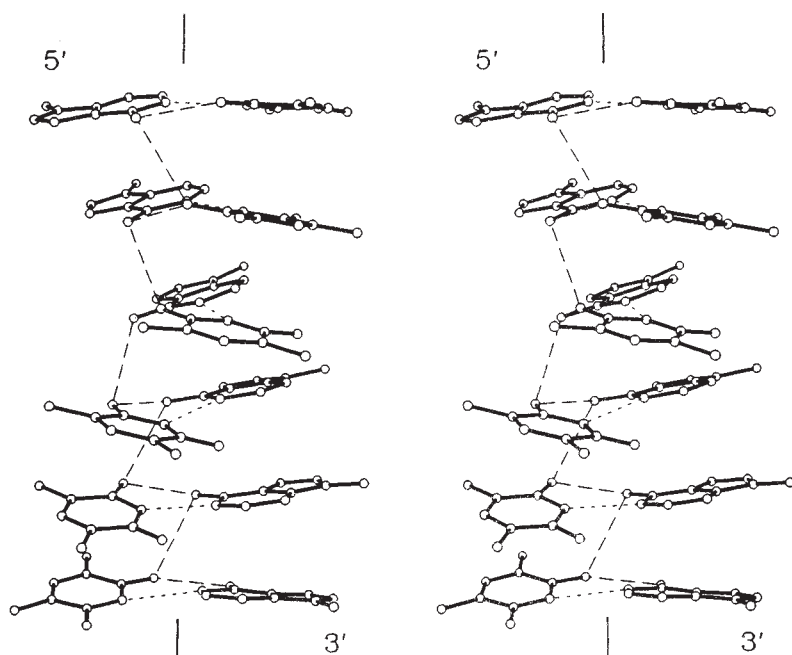


Fig. 4 The oligo(dA)-oligo(dT) stretch, illustrating the bases and best helix axis. The 5' and 3' ends of the adenine strand are labelled accordingly. The Watson-Crick N1-adenine/N3-thymine distances are: A4 to T21, 2.6 Å; A5 to T20, 2.8 Å; A6 to T19, 2.7 Å; A7 to T18, 2.7 Å; A8 to T17, 2.7 Å; and A9 to T16, 2.7 Å. The Watson-Crick N6-adenine/O4-thymine distances are: A4 to T21, 2.7 Å; A5 to T20, 3.1 Å; A6 to T19, 3.1 Å; A7 to T18, 2.7 Å; A8 to T17, 2.8 Å; and A9 to T16, 2.7 Å. The cross-strand diagonal N6-adenine/O4-thymine distances are: A4 to T20, 3.2 Å; A5 to T19, 3.1 Å; A6 to T18, 3.4 Å; A7 to T17, 3.2 Å; and A8 to T16, 3.3 Å. Although the positions of the hydrogen atoms cannot be located at the resolution of this crystal study, when they are put into idealised positions, the resulting cross-strand diagonal bonds fit the criteria for the long component of the bifurcated hydrogen bonds as described by Taylor *et al.*³². At the present time, we have no explanation for the one longish bond in the middle (A6 to T18). It ought to be remembered that the energy of the hydrogen bond, being largely electrostatic in nature, does not fall off sharply with distance. Indeed, energy calculations have predicted cross-strand diagonal interactions (M. Levitt, personal communication).

of the structure. However, the effect of this compact stacking arrangement on the sugar-phosphate chains is to narrow the width of the minor groove (see Fig. 3). If measured by the distance between phosphorus atoms on opposite strands staggered by three base pairs, the minor groove in the A·T tract is on average 9.5 ± 0.4 Å wide.

It has been suggested that a high degree of propeller twist (and consequent narrowing of the minor groove) might be stabilized by a minor groove spine of hydration, whereby solvent molecules are highly ordered by tetrahedral coordination between cross-strand pyrimidine O2 and purine N3 atoms one base pair apart^{27,29}. Although we find solvents in the major groove and near the sugar-phosphate chains, we have not located ordered solvent atoms in the minor groove. But, owing to the limited resolution of this study, in which only 27 solvents have been unambiguously located, we cannot comment definitively on the state of minor groove solvation.

Roll, tilt and junctions

The path of a helix can be described by calculating the orientation of the best mean plane of the base pairs, either with respect to a successive base pair (defining the orientation of a base step), or with respect to a helix axis. The latter calculation is heavily dependent on the choice of the best helix axis, which can be particularly difficult if the path of the helix is changing. Table 1 shows the helical parameters describing the orientation of the base pairs; the definitions of these parameters are in the table legend. With the exception of the last base step (A8, A9-T16, T17), discussed in the legend to Fig. 6, the roll and tilt angles for the AA-TT base steps are close to zero. The extreme

variation of the base steps at the ends of the structure make the minimal variation within the AA-TT base steps even more striking. The near-zero values for the roll, tilt and slide of the AA-TT base steps indicate that each A·T base pair is not changing position with respect to its neighbours, the base pairs being parallel to each other.

Because of the near-uniform geometry of the A·T tract, a best-helix axis was calculated for this region alone and used as a reference to examine the path of the rest of the helix. Figure 6 shows the direction of the normals to the base pairs, plotted as projections of vectors relative to a common origin. Two features about the direction of the helix are obvious from this diagram. First, the projection of vectors for the A·T base pairs (especially for the vectors 4 to 8, corresponding to the first five A·T base pairs) are short and clustered together about the origin, indicating that the base pairs are not only parallel to each other, but perpendicular to the helix axis. Second, the overall direction of the vectors changes from right to left in the diagram, indicating an overall change in the helix direction. These changes occur most abruptly at two places: between vector 3 and vector 4 (that is, between base pairs C3-G22 and A4-T21) and between vector 10 and vector 11 (between base pairs G10-C15 and C11-G14). Thus, the overall picture from the vector diagram is a straight bit of oligo(dA)-oligo(dT) with two junctions having additive effects on the change in the helix direction.

The nature of these particular junctions are worth describing in light of the junction-based theories for DNA bending. They can be explained simply as a combination of base-stacking and propeller twist. In B-DNA, at a 5'-pyrimidine-purine-3' base step, the propeller twist and larger size of purine bases causes

Fig. 5 Stereo views of a typical AA-TT base step. The two base pairs, A5-T20 (in black) and A6-T19 (in white) are viewed down the helix axis (illustrated by a black circle). Compare this base stacking arrangement to that found in poly(dG)-poly(dC), where the guanines stack in a staggered fashion (six-membered ring over five-membered ring), and the cytosines show no overlap⁴⁶.

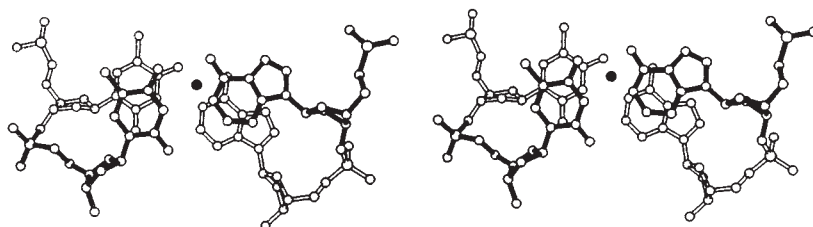
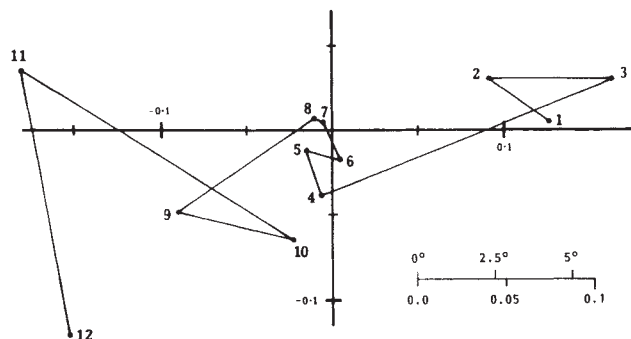


Fig. 6 Polar diagram of normals to the mean plane of the base pairs, projected onto a plane perpendicular to the best helix axis of the AT region. Each point specifies the end of a unit vector drawn perpendicular to the mean plane of the base pair, all vectors being drawn from the same origin. Vector 1 is the normal to the base pair C1-G24; vector 2 is the normal to the base pair G2-C23, and so on. The length of each vector from the origin is the arc sin of the angle between the mean plane of each base pair and the perpendicular to the best helix axis⁴⁷. The distance between any two vector end points is the arc sin of the angle between the mean planes through each base pair (thus, an angle related to a combination of the roll and tilt angles of the base steps). A similar plot constructed for the CGCGAATTCGCG dodecamer shows a similar relationship between the points at the end of the structure; this is due to the similarity in the crystal packing. But the distribution of the vectors in the central region is quite different, indicating a difference in the distribution of bending. The overall bend in the CGCGAATTCGCG dodecamer is delocalized over several base steps, as opposed to the abrupt changes seen for our structure. Although vector 9 (the last A-T base pair, A9-T16) is much larger and separate from the other AT vectors, vector 10 (the G-C base pair at the 3' end of the adenine stack) is stacked more with the rest of the A-T base pair. This presumably is an end-effect arising because of the conflicting demands of the G-C base pair trying to maximize purine-purine stacking on the A-T base pair, yet unable to undergo the same degree of high propeller twist as the A-T base pairs.



steric clashes between the cross-strand purine bases in the minor groove^{36,37}. When the second base pair has a larger degree of propeller twist (effectively pushing the minor groove towards the 5' side), the normal pyrimidine-purine minor groove steric clash is exaggerated by the difference in the propeller twist causing a larger than normal roll away from the minor groove at that step, as seen by the 10° roll at the C3, A4-T21, G22 base step. At a 5'-purine-pyrimidine-3' step, the purines now clash in the major groove. Again, the effects of the differential propeller twist will affect this. The situation at the 3' end of the adenine stack of the structure is made more complicated by the fact that the junction is not a purine-pyrimidine. If, however, one assumes the G10-C15 base pair to be an honorary A-T base pair (as described in the legend to Fig. 6), then the purine-pyrimidine clash/propeller twist clash occurs in the major groove of the G10, C11-G14, C15 base step. Again, there is a large roll angle, but it is in the opposite direction (-10°).

Consequences of structure

The fact that the properties of poly(dA)·poly(dT) in solution and in fibres show a conformational rigidity strongly indicates

that the crystal structure of the oligo(dA)·oligo(dT) tract can validly represent a structure of poly(dA)·poly(dT) (see Fig. 7). Indeed, the distinctive features which have emerged (additional hydrogen bonds and good base stacking) are such as would confer a rigidity over and above that expected from base pairs with two hydrogen bonds. It now becomes obvious why the homopolymer will not wrap into nucleosome cores, and why short runs take up positions of minimum curvature: the structure resists bending because it would destroy the good base-stacking and additional hydrogen bonds. The decreased helical repeat, reduced axial rise, and narrow minor groove in relation to standard B-type DNA are just natural consequences of the compact base-stacking arrangement.

How does this structure relate to the bending of DNA by phased runs of oligo(dA)·oligo(dT)? There must be something else about the poly(dA)·poly(dT) structure other than resistance to bending, or else permanent bending could be accomplished by runs of oligo(dG)·oligo(dC). The perpendicularity of the mean plane of the A-T base pairs to the helix axis, with roll and tilt angles near zero, as well as the high degree of propeller twist within the base pairs, must be significant.

Fig. 7 Stereo view of a CPK model of the homopolymer poly(dA)·poly(dT), created by extension of the double helix from the middle four A-T base pairs. The high degree of propeller twist causes the double helix to have a deep, narrow minor groove and a wide, shallow major groove. This contrasts with the model for homopolymer poly(dG)·poly(dC), built by extension from the crystal structure of the GGGGCCCC octamer, which has a wide, shallow minor groove due to the displacement of the G-C base pairs into the major groove to increase purine-purine stacking⁴⁶. Although the models from the recent solution and fibre studies of poly(dA)·poly(dT) agree with the crystal structure in terms of the gross features of narrow minor groove and non-coplanarity of the base pairs, they disagree in terms of the deconvolution of the non-coplanarity into propeller twist, as well as an absolute tilt of about 6° of the base pair with respect to the helix axis⁴⁸⁻⁵⁰. This base pair tilt is not seen in the crystal structure; the average mean plane through an AT base pair is perpendicular to a best helix axis, although the base pairs do exhibit a variable degree of buckle. It is not surprising that these models differ from the crystal structure in this respect due to the inherent difficulties involved in fitting models to cylindrically averaged data at relatively low resolution.



In the literature, there have been several theories of bending which differ in their predictions as to where the actual changes in the direction of the helix axis occur. The wedge model in its latest form proposes that the AT stretch itself is bent due to a combination of roll and tilt at each and every AA-TT step. This large, non-coplanar wedge, estimated to be around 9°, would cause a change in the helix direction with respect to adjoining pieces of DNA (ref. 38). We have shown that the AA-TT steps do not have roll or tilt components, and hence the oligo(dA)·oligo(dT) stretch cannot cause a change in orientation of this sort. This rules out the wedge model and implies that any bending of the helix axis must occur in the regions between the runs of oligo(dA)·oligo(dT).

The junction model suggests that the bending occurs abruptly at the ends of the AT stretch due to some specialised features of that region^{12,19}. In fact, we suggest that the larger-than-usual roll angles at the base steps at the ends of the oligo(dA)·oligo(dT) occur partially as a result of the high degree of propeller twist from the structure. Much of the contribution to the overall bending could come from the junctions if, as here, the roll angles at the ends of the A·T stretch had opposite directions with respect to the minor and major grooves, thus causing an additive effect when placed half a helical turn apart. There is a small tilt component to the bending (as seen in Table 1), but it is not as dramatic as the roll component.

We note, however, that it is not in general necessary to have all of the change concentrated at the junctions; it could be more widely distributed over the regions between the AT tract. To give a permanent bend, there need only be a net roll accumulated in the intervening regions, as follows from the considerations given by Calladine and Drew³⁹. This can be understood most simply by considering two successive half turns of a uniform double helix, which has a fixed roll angle between base

pairs (as in Fig. 2 of ref. 40). After one full helical turn, the base pairs will be parallel to each other. However, each half turn has a net roll angle between its first and last base pairs (obtained by vectorial summation of the individual contributions). Because of the helical twist between the two halves, they have resultants of opposite sign and so cancel when adjacent; therefore, the helix axis remains unchanged. But if one half turn is replaced by an AT stretch of zero net roll, then the net roll contributed by the remaining half is not cancelled, and the direction of the helix axis will change.

Without further direct studies of structures including the intervening regions, it is difficult to decide between an abrupt change at the junctions, a gradual change over the intervening region, or a combination of the two. Whatever the case, it is the contrast between the special geometry of the oligo(dA)·oligo(dT) tract and that of other DNA sequences that is the basis of the bending.

We thank Horace Drew, Bill Hunter, Michael Levitt, Maxine McCall and Daniela Rhodes for advice and support throughout the whole of this project. In addition, we would like to acknowledge: Terry Smith, for synthesis of oligonucleotides; Dan Brown, for help with purification; Phil Evans and Judd Fermi, for help with refinement; and Chris Calladine and Andrew Travers for comments on the manuscript. This paper is dedicated to Max Perutz in honour of the 40th anniversary of the MRC Laboratory of Molecular Biology. H.C.M.N. is a Fellow of The Jane Coffin Childs Memorial Fund for Medical Research. This investigation has been aided by a grant from the Jane Coffin Childs Memorial Fund for Medical Research. This work was presented at the Gordon Research Conference on Nucleic Acids on 15 June 1987. *Note added in proof:* Bifurcated hydrogen bonds have now been recognized in the dodecamer d (CGCAAATTGCG) crystallized with and without the drug distamycin⁵¹.

Received 30 August; accepted 19 October 1987.

- Peck, L. J. & Wang, J. C. *Nature* **292**, 375-378 (1981).
- Rhodes, D. & Klug, A. *Nature* **292**, 378-380 (1981).
- Arnot, S. & Selsing, E. *J. molec. Biol.* **88**, 509-521 (1974).
- Leslie, A. G. W., Arnot, S., Chandrasekaran, R. & Ratliff, R. L. *J. molec. Biol.* **143**, 49-72 (1980).
- Rhodes, D. *Nucleic Acids Res.* **6**, 1805-1816 (1979).
- Simpson, R. T. & Kunzler, P. *Nucleic Acids Res.* **6**, 1387-1415 (1979).
- Kunkel, G. R. & Martinson, H. G. *Nucleic Acids Res.* **9**, 6869-6888 (1981).
- Prunell, A. *EMBO J.* **1**, 173-179 (1982).
- Struhl, K. *Proc. natn. Acad. Sci. U.S.A.* **82**, 8419-8432 (1985).
- Satchwell, S. C., Drew, H. R. & Travers, A. A. *J. molec. Biol.* **191**, 659-675 (1986).
- Travers, A. & Klug, A. *Phil. Trans. R. Soc.* (in the press).
- Koo, H.-S., Wu, H.-M. & Crothers, D. M. *Nature* **320**, 501-506 (1986).
- Hagerman, P. J. *Nature* **321**, 449-450 (1986).
- Diekmann, S. *Nucleic Acids Res.* **15**, 247-265 (1987).
- Hagerman, P. J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4632-4636 (1984).
- Levene, S. D., Wu, H.-M. & Crothers, D. M. *Biochemistry* **25**, 3988-3995 (1986).
- Griffith, J., Bleyman, M., Rauch, C. A., Kitchin, P. A. & Englund, P. T. *Cell* **46**, 717-724 (1986).
- Marini, J. C., Levene, S. D., Crothers, D. M. & Englund, P. T. *Proc. natn. Acad. Sci. U.S.A.* **79**, 7664-7668 (1982).
- Wu, H.-M. & Crothers, D. M. *Nature* **308**, 509-513 (1984).
- Bossi, L. & Smith, D. M. *Cell* **39**, 643-652 (1984).
- Snyder, M., Buchman, A. R. & David, R. W. *Nature* **324**, 87-89 (1986).
- Zahn, K. & Blattner, F. R. *Science* **236**, 416-427 (1987).
- Stenzel, T. T., Patel, P. & Bastia, D. *Cell* **49**, 709-717 (1987).
- Ryder, K., Silver, S., DeLucia, A. L., Fanning, E. & Tegtmeyer, P. *Cell* **44**, 719-725 (1986).
- Wing, R. *et al. Nature* **287**, 755-758 (1980).
- Drew, H. R. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 2178-2183 (1981).
- Dickerson, R. E., Kopka, M. L. & Pjura, P. in *Biological Macromolecules and Assemblies: Nucleic Acids and Interactive Proteins* (eds Jurnak, F. A. & McPherson, A.) **2**, 38-126 (Wiley, 1984).
- Hunter, W. N. *et al. J. biol. Chem.* **262**, 9962-9970 (1987).
- Fratini, A. V., Kopka, M. L., Drew, H. R. & Dickerson, R. E. *J. biol. Chem.* **257**, 14686-14707 (1982).
- Dickerson, R. E. & Drew, H. R. *J. molec. Biol.* **149**, 761-786 (1981).
- Donohue, J. in *Structural Chemistry and Molecular Biology* (eds Rich, A. & Davidson, N.) **1**, 450-456 (Freeman, San Francisco, 1968).
- Taylor, R., Kennard, O. & Versichel, W. *J. Am. chem. Soc.* **106**, 244-248 (1984).
- Baker, E. N. & Hubbard, R. E. *Prog. Biophys. molec. Biol.* **44**, 98-179 (1984).
- Jack, A., Ladner, J. E. & Klug, A. *J. molec. Biol.* **108**, 619-649 (1976).
- Prive, G. G. *et al. Science* **238**, 498-504 (1987).
- Calladine, C. R. *J. molec. Biol.* **161**, 343-352 (1982).
- Dickerson, R. E. *J. molec. Biol.* **166**, 419-441 (1983).
- Ulanovsky, L. E. & Trifonov, E. N. *Nature* **326**, 720-722 (1987).
- Calladine, C. R. & Drew, H. R. *J. molec. Biol.* **192**, 907-918 (1986).
- Calladine, C. R. & Drew, H. R. *J. molec. Biol.* **178**, 773-782 (1984).
- Sussman, J. L., Holbrook, S. R., Church, G. M. & Kim, S. H. *Acta crystallogr. sect. A* **33**, 800-804 (1977).
- Jones, T. A. *J. appl. Crystallogr.* **11**, 268-272 (1978).
- Hendrickson, W. A. & Konner, J. H. in *Biomolecular Structure and Conformation, Function and Evolution* (ed. Srinivasan, R.) **1**, 43-57 (Pergamon, Oxford, 1980).
- Westhof, E., Dumas, P. & Moras, D. *J. molec. Biol.* **184**, 119-145 (1985).
- Arnot, S., Chandrasekaran, R., Hall, I. H. & Puigjaner, L. C. *Nucleic Acids Res.* **11**, 4141-4155 (1983).
- McCall, M., Brown, T. & Kennard, O. *J. molec. Biol.* **183**, 385-396 (1985).
- Dickerson, R. E., Kopka, M. L. & Pjura, P. *Proc. natn. Acad. Sci. U.S.A.* **80**, 7099-7103 (1983).
- Edmondson, S. P. & Johnson, W. C. Jr *Biopolymers* **24**, 825-843 (1985).
- Alexeev, D. G., Lipanov, A. A. & Skuratovskii, I. Y. *Nature* **325**, 821-823 (1987).
- Park, H.-S., Arnot, S., Chandrasekaran, R., Millane, R. P. & Campagnari, F. *J. molec. Biol.* **197**, 513-523 (1987).
- Coll, M., Frederick, C. A., Wang, A. H.-J. & Rich, A. *Proc. natn. Acad. Sci. U.S.A.* (in the press).

Discovery of hard X-ray emission from supernova 1987A

R. Sunyaev*, A. Kaniovsky*, V. Efremov*, M. Gilfanov*,
E. Churazov*, S. Grebenev*, A. Kuznetsov*,
A. Melioranskiy*, N. Yamburenko*, S. Yunin*,
D. Stepanov*, I. Chulkov*, N. Pappe*, M. Boyarskiy*,
E. Gavrilova*, V. Loznikov*, A. Prudkoglyad*,
V. Rodin*, C. Reppin†, W. Pietsch†, J. Engelhauser‡,
J. Trümper†, W. Voges†, E. Kendziorra‡, M. Bezler†,
R. Staubert†, A. C. Brinkman§, J. Heise§, W. A. Mels§,
R. Jager§, G. K. Skinner||, O. Al-Emam||,
T. G. Patterson|| & A. P. Willmore||

* Space Research Institute, USSR Academy of Sciences, Moscow, USSR

† Max-Planck-Institut für Physik und Astrophysik, Institut für Extraterrestrische Physik, 8046 Garching, FRG

‡ Astronomisches Institut der Universität Tübingen, 7400 Tübingen, FRG

§ Space Research Laboratory, Utrecht, The Netherlands

|| Department of Space Research, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

We report the discovery of hard X-rays from the region of the supernova SN1987A in the Large Magellanic Cloud. The observations were made from the Mir-Kvant observatory 'Röntgen'. Hard X-rays were first observed on 10 August 1987 and SN1987A then became the main target of the observatory. Up to 15 September 1987, a total of 115 pointings on 21 days were made. The flux showed little variation during the observation period. The measured spectrum extends from 20 keV to 300 keV and is extremely hard, having a photon power law index of ~ 1.4 . At low energies the spectrum becomes even flatter and there is indication of a cutoff between 10 and 25 keV. The luminosity over the energy range 20–300 keV is $\sim 2 \times 10^{38} \text{ erg s}^{-1}$ (assuming a distance of 55 kpc). The error box for the hard source has a (2σ) radius of 10 arc min and contains SN1987A. The results have important implications for current models of SN1987A.

Before the discovery of the hard X-ray flux, the Kvant-Röntgen Observatory observed SN1987A on 8 June, 16 July and 1 August 1987. The hard X-ray source was first detected with the high energy X-ray experiment (HEXE) instrument. This detector consists of four NaI/CsI phoswich scintillators with a total geometric area of 800 cm^2 , with an energy range 15–250 keV (see ref. 1). The field of view is defined by two hexagonal collimators with 1.6° full-width-at-half-maximum (FWHM) response, which can be tilted individually by 2.3° for background measurements at 2-minute intervals. The crosses in Fig. 1 show the spectrum obtained during the 10–21 August 1987 period using this collimator rocking mode for 9,600 s. The conversion from the phoswich pulse height spectrum has been performed using a detector response matrix based on laboratory calibrations. Crab nebula observations performed with HEXE in September 1987 were analysed in the same way and gave a satisfactory fit. Although the data are still preliminary, we estimate that any systematic errors are less than 30%.

To locate the hard X-ray source we first employed the Coded Mask Imaging Spectrometer (TTM) telescope (ref. 2), which images a field of view of 7.8° by 7.8° FWHM, with an angular resolution of 1.8 arc min. It operates in the energy range 2–32 keV, although the efficiency of the xenon-filled proportional counter detector is greatest in the lower part of this range.

Figure 2 shows part of an image obtained by combining 40,000 s of data from 50 pointings in the period 10–25 August. LMC X-1, X-2, X-3 and X-4 are immediately apparent. The first three of these sources are detected in most individual

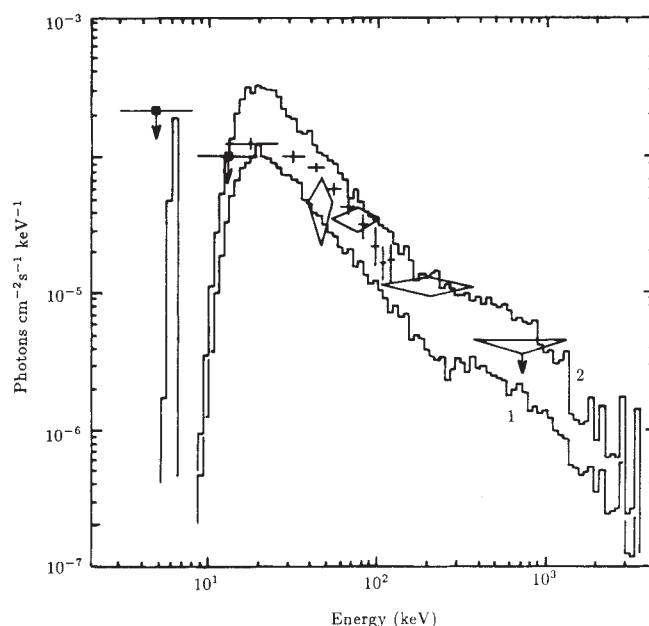


Fig. 1 Preliminary energy spectrum of the hard X-rays. Squares, 3σ upper limits from TTM; crosses, detections with HEXE (1σ error bars); diamonds, detections and 3σ upper limits obtained with Pulsar X-1. The histograms show the results of Monte Carlo simulations by Grebenev and Sunyaev¹⁰, assuming $0.1 M_\odot$ of ^{56}Co , an expanding envelope of $16 M_\odot$ with metallicity 1/3 solar, and mean expansion velocity $4,150 \text{ km s}^{-1}$, 180 days (1) and 210 days (2) after the explosion. The histogram bins in the region of the Fe fluorescence line are 0.5 keV wide.

pointings. Figure 3 shows an expansion of the part of the image surrounding the position of SN1987A. The data show that at low energies LMC X-1 and the 50-ms pulsar 0540–693 are the only strong sources close to SN1987A, and demonstrate that if the flux at higher energies does not come from either of these objects then it must be from a source with a very unusual spectrum.

It has not yet been possible to confirm the response of the TTM instrument by observations of the Crab nebula, but 0540–693 provides a convenient calibration source. Taking its spectrum to be $0.006 E^{-1.8} \text{ photons cm}^{-2} \text{ s}^{-1} \text{ keV}^{-1}$ (where E is the photon energy in keV) (refs 3, 4) we find a 3σ upper limit of 5×10^{-4} times the intensity of the Crab nebula for any emission from SN1987A with a similar spectrum.

As the TTM data do not allow a direct confirmation of the origin of the high-energy flux and because of the 1.6° width of the FWHM of the HEXE collimators, most of the 115 pointings (each of which was typically of 20 min duration) were made with offsets of up to 0.8° from the position of SN1987A to allow discrimination between SN1987A and other sources. Figure 4 shows an analysis of measurements from 54 pointings in the form of probability contours for the source position. The supernova lies close to the centre of the 1σ contour, whereas LMC X-1 and 0540–693 are well outside the 3σ contour.

To extend the spectral information to higher energies, the Pulsar X-1 instrument was used. This instrument includes four 20-cm-diameter NaI/CsI phoswich detectors which are $3 \times 3 \text{ cm}$ thick and are optimized for the energy range 50–1,000 keV. As the instrument has a 3° FWHM fixed collimator, background observations are made by rocking the entire space-station complex by 12° , spending 4 min on-source and 4 min off-source. This mode is inconvenient for the other experiments, so relatively few observations have been made (a total of 5,000 s on-source), but 5σ detections, shown in Fig. 1, were obtained in two broad bands extending from 50 keV up to 300 keV.

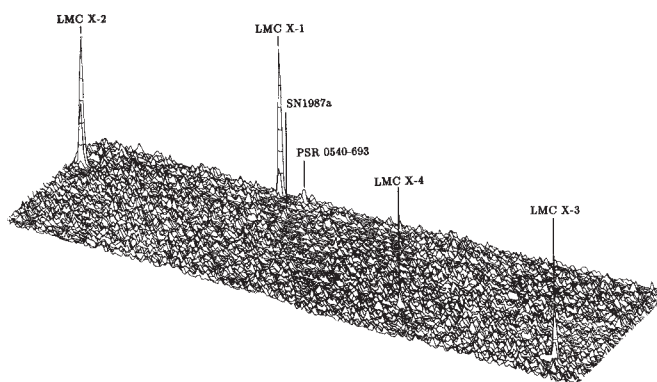


Fig. 2 A 2.5° by 8.5° slice of an image in the energy range 2–32 keV obtained with the TTM instrument during the HEXE observations. There is no peak at the position of SN1987A. Vertical height indicates significance of detection, not absolute intensity.

Also shown in Fig. 1 are upper limits in the 2–10 and 10–20 keV bands, assuming a flat spectrum, from the TTM data. The limits are consistent with the HEXE data and also with that from Ginga⁵ (but see note added in proof). Although the data are preliminary it is clear that the spectrum is relatively flat (much flatter than the $\sim E^{-2}$ spectrum of the Crab nebula, for example) and that it must change its slope below 25 keV.

Unfortunately the observations before 10 August were too short for a sensitive detection, but the preliminary analysis shows that the source was at least not brighter on these dates. The HEXE data obtained during the period 10 August to 15 September indicate that the source intensity remained approximately constant. Any flux increase during this period cannot have been by more than 30%.

The hard X-ray flux is unlikely to be the result of thermal emission from a hot plasma compressed and heated by a shock wave; even if the ions were heated to the required temperature of $\sim 10^9$ K they would transfer energy to the electrons only inefficiently. But such a shock must exist, and our measurements can be used to find an upper limit for the emission measure of the hot gas. Using formulae for the bremsstrahlung and recombination spectral emissivity of a compressed gas, and assuming a distance of 55 kpc, we find $N_e^2 V < 3 \times 10^{59} \text{ cm}^{-3}$ (where N_e is the electron number density and V is the volume) for electron temperatures T_e in the broad range $3 < kT_e < 300 \text{ keV}$ (where k is the Boltzmann constant). In fact, again using standard formulae (see, for example, ref. 6), we expect that 0.5 yr after the explosion the electron temperature will be $kT_e = 10 N_5^{0.4} v_{10,000}^{0.8} \text{ keV}$, where N_5 is the post-shock electron density in units of 10^5 cm^{-3} and $v_{10,000}$ is the front velocity in units of $10,000 \text{ km s}^{-1}$.

If one assumes that the density distribution in the external part of the supernova shell is given by $\rho \propto R^{-7}$ and if $v \propto R$, it follows that for ambient densities in the relevant range the present velocity of the shock front is $v_{\text{front}} = 17,000 \text{ km s}^{-1}$ or higher⁷, but we make here the moderate assumption $v_{\text{front}} = 10,000 \text{ km s}^{-1}$. The outer radius of the shocked region is equal to $v_{\text{front}} t$, where t is the time since the explosion, and taking into account the fourfold compression of the gas in the shock^{8,9}, it is easy to estimate the volume of compressed gas using only the condition of mass conservation as $(\pi/3)(v_{\text{front}} t)^3$. Using this volume and the above upper limit on the emission measure, we find $N_e < 2.7 \times 10^5 \text{ cm}^{-3}$. The ambient gas density will be 4 times lower ($< 7 \times 10^4 \text{ cm}^{-3}$). The bremsstrahlung cooling time for such a compressed gas density exceeds tens of years and the time for Compton cooling due to interaction of hot electrons with optical and infrared photons exceeds 30 yr.

From the above limit to the density of the ambient plasma one can estimate the mass loss rate of the pre-supernova to be

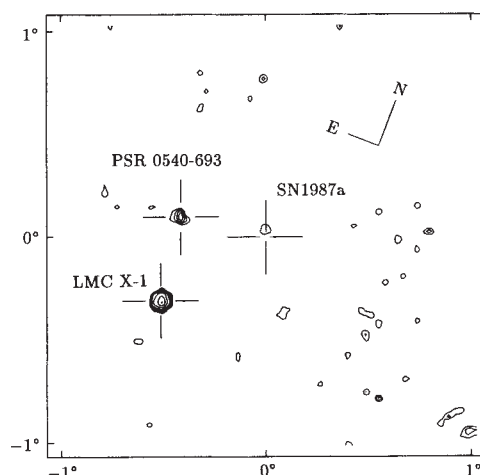


Fig. 3 Contour map of the 2° square region of the image in Fig. 2 surrounding SN1987A. Axes show offsets from the position of the supernova. Apart from LMC X-1 and 0540-693, the levels seen are consistent with the expected noise level. Contours are at 3, 4, 5, 6, 7, 8, 10, 20, 40 and 60×10^{-1} uncorrected counts per second.

$\dot{M} < 5 \times 10^{-5} (v_w/100 \text{ km s}^{-1}) M_\odot \text{ yr}^{-1}$, where v_w is the stellar wind velocity of the progenitor and $M_\odot$ is the mass of the Sun. We note that this estimate applies to a time of $\sim 0.5 \text{ yr} \times (v_{\text{front}}/v_w)$, or ~ 20 – 100 yr before the explosion.

It is most probable that the observed hard X-rays arise from radiation generated within the core of the supernova which is starting to percolate out. Possible origins that can be considered for the photons include: (1) gamma radiation from radioactive decay of ^{56}Co (refs 10, 11); (2) radiation from a young pulsar (period 0.5–100 ms); (3) accretion onto a collapsed object, either a single star or a member of binary system; and (4) thermal radiation from a hot neutron star.

For a homologous, spherically symmetrical, uniform-density model the Thomson optical depth is $\sim 5 M_{16} \hat{v}_{4,000}^{-2} t_y^{-2}$, where M_{16} is the mass normalized to $16 M_\odot$, $\hat{v}_{4,000}$ is the average velocity in units of $4,000 \text{ km s}^{-1}$ and t_y is the time in years after the explosion. Thus the supernova is expected to be still opaque to X- and gamma rays. If the original photon energy is high enough, photons can undergo a number of Compton scatterings before their energy is degraded to the point where photoelectric absorption becomes probable and after ~ 6 months the envelope is expected to become thin enough that a small proportion of photons escape before they can be absorbed. The fraction of the energy that escapes is still $< 10^{-3}$ and so we can exclude cases (3) and (4) as the luminosity of the internal source would have to be greatly in excess of (1,000 times) the Eddington limit for an object of a few solar masses. Accretion onto a collapsed object would be halted and it is difficult to believe that the luminosity of a cooling neutron star could be so high.

Figure 1 shows the results of Monte Carlo simulations by Grebenev and Sunyaev¹² for case (1), radio-active decay, and fig. 5 shows the corresponding simulations for case (2), a buried pulsar. Xu *et al.*¹³ and others have reached similar conclusions. Given that the precise time at which a given flux level is expected to be reached is somewhat uncertain (depending on the mass, expansion velocity and proportion of heavy elements), there is satisfactory agreement between the spectrum and either of these models. The relative constancy of the flux during the observations is, however, unexpected.

On the basis of the radioactive decay model, the flux is expected to rise to a peak (at ~ 300 days with the parameters used in Fig. 1) and then decline as the ^{56}Co becomes exhausted. The absence of any strong increase in the observed flux may be

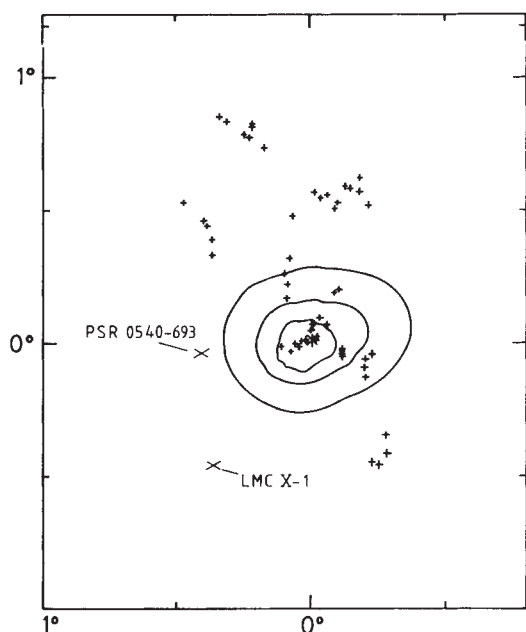


Fig. 4 The position of the source of high-energy flux obtained by analysing 54 HEXE observations at the offset positions shown (crosses). Contours are at 1σ , 2σ and 3σ confidence. The positions of LMC X-1 and 0540-69.3 are excluded at the 3σ level.

explained if the peak has been reached earlier than predicted by current models.

One reason for this discrepancy may be that the assumption of spherical symmetry is not valid. Inhomogeneities may have developed because of the inherent Rayleigh-Taylor instability of the envelope, or the structure may be toroidal. Simulations of these situations¹² have shown that the X-ray emission may then appear earlier and be more constant.

In the case of the pulsar model the flux is expected to rise continuously. The standard model assumed in Fig. 5 (a $16 M_{\odot}$ spherically symmetric shell) predicts an increase by a factor of ~ 3 per month throughout the first year, whereas the observational limit is $\sim 30\%$ per month. Again it is possible that there may not be spherical symmetry.

Whatever the energy source, the upper limits placed on the low-energy flux by the TTM observations provide support for the models in which high-energy flux is down-Comptonized as it diffuses out through an absorbing envelope. Further analysis of the TTM and HEXE data should be able to show whether photoelectric absorption is really occurring.

Observations are continuing and will provide a means of distinguishing between cases (1) and (2). In the case of a pulsar (2) the flux is expected to increase, whereas if the energy source is radioactive decay (1) then a decline should set in within <1 yr of the supernova explosion. In the latter case a lower-luminosity pulsar or accreting neutron star or black hole (either single or in a close binary) may eventually be revealed and could be

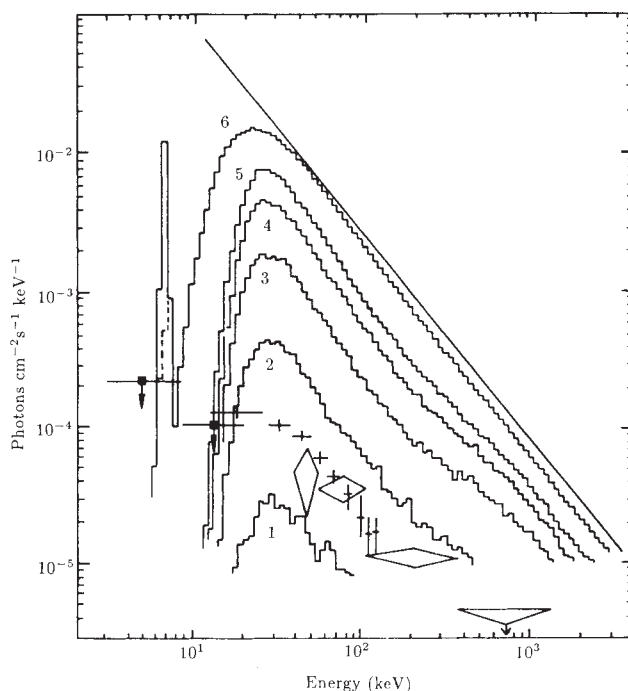


Fig. 5 As Fig. 1, except that the simulations assume that the central energy source is a pulsar with a luminosity in the energy range 1–1,000 keV of 8×10^{41} erg s⁻¹ at 1–1,000 keV and a spectrum similar to that of the Vela pulsar (photon index 1.5). The predictions are for 180, 240, 300, 360, 420 and 720 days after the explosion (labelled 1–6 respectively).

detected with the present instrumentation if the X-ray luminosity is of the order of 10^{37} – 10^{38} erg s⁻¹.

Note added in proof: The data presented here are in reasonable agreement with those obtained with the Ginga satellite at energies >12 keV, but upper limits placed on the low-energy emission with the TTM instrument are inconsistent with the existence of a low energy component at the level seen in the Ginga data. The observations are not strictly simultaneous but between Ginga observations a few days on either side of the TTM ones the reported flux increased by only 20%. It is emphasized that this analysis of our data is preliminary and in particular that there has been no Crab calibration of the TTM sensitivity. Nevertheless we think it unlikely that using 0540-69.3 as an intensity calibration and assuming the Einstein spectrum introduces significant errors. A source with the spectrum found by Ginga ought to appear in the TTM images at $>6\sigma$ with an intensity 1.6 times that of 0540-69.3. The only effect that we have been able to find which would tend to reduce the discrepancy is a small negative bias to the TTM gap in the region of SN1987A but hardly affecting 0540-69.3. The effect amounts to 10% of the upper limits quoted. We note that the limits which we deduce for the density of the medium surrounding SN1987A would remain valuable even with X-ray limits several times higher.

Received 7 October; accepted 27 October 1987.

1. Reppin, C., Pietsch, W., Trümper, J., Kendziorra, E. & Staubert, R. *Proc. Workshop Non-thermal and Very High Temperature Phenomena in X-ray Astronomy, Rome, 1983* (eds Perola, G. C. & Salvati, M.) 279–282 (Istituto Astronomica Roma, Rome, 1985).
2. Brinkman, A. C., Dam, J., Mels, W. A., Skinner, G. K. & Willmore, A. P. *Proc. Workshop Non-thermal and Very High Temperature Phenomena in X-ray Astronomy, Rome, 1983* (eds Perola, G. C. & Salvati, M.) 263–270 (Istituto Astronomica Roma, Rome, 1985).
3. Seward, F. D., Harnden, F. R. & Helfand, D. J. *Astrophys. J.* **287**, L19–L22 (1984).
4. Clark, D. H. *et al. Astrophys. J.* **255**, 440–446 (1982).

5. Makino, F. *IAU Circ. No. 4447* (1987).
6. Spitzer, L. *Physics of Fully Ionized Gases* (Interscience, New York, 1956).
7. Gifanov, M., Sunyaev, R. A. & Churazov, E. *Soviet Astr. Lett.* **14**, 1–6 (1988).
8. Zeldovich, Ya. B. & Raizer, Yu. P. *Physics of Shock Waves* (Academic, New York, 1966).
9. Chevalier, R. A. *Rev. Astr. Astrophys.* **15**, 175–196 (1977).
10. Clayton, D. D., Colgate, S. A. & Fishman, G. J. *Astrophys. J.* **155**, 75–82 (1969).
11. Colgate, S. A. & McKee, C. *Astrophys. J.* **157**, 623–643 (1969).
12. Grebenev, S. A. & Sunyaev, R. A. *Soviet Astr. Lett.* **13**, 945–963 (1987).
13. Xu, Y., Sutherland, P., McCray, R. & Ross, R. R. *Astrophys. J. Lett.* (in the press).

Discovery of an unusual hard X-ray source in the region of supernova 1987A

T. Dotani*, K. Hayashida*, H. Inoue*, M. Itoh*, K. Koyama*, F. Makino*, K. Mitsuda*, T. Murakami*, M. Oda*, Y. Ogawara*, S. Takano*, Y. Tanaka*, A. Yoshida*, K. Makishima†, T. Ohashi†, N. Kawai‡, M. Matsuoka‡, R. Hoshi§, S. Hayakawa||, T. Kii||, H. Kunieda||, F. Nagase||, Y. Tawara||, I. Hatsukade†, S. Kitamoto¶, S. Miyamoto¶, H. Tsunemi¶, K. Yamashita¶, M. Nakagawa¶, M. Yamauchi¶, M. J. L. Turner**, K. A. Pounds**, H. D. Thomas**, G. C. Stewart**, A. M. Cruise††, B. E. Patchett†† & D. H. Reading††

* Institute of Space and Astronautical Science, Komaba, Meguro-ku, Tokyo, Japan

† Department of Physics, University of Tokyo, Hongo, Bunkyo-ku, Tokyo, Japan

‡ Institute of Physical and Chemical Research, Wako, Saitama, Japan

§ Department of Physics, Rikkyo University, Nishi-Ikebukuro, Toshima-ku, Tokyo, Japan

|| Department of Astrophysics, Nagoya University, Chikusa-ku, Nagoya, Japan

¶ Department of Physics, Osaka University, Toyonaka, Japan

¶ Department of Physics, Osaka City University, Sugimoto-cho, Sumiyoshi-ku, Osaka, Japan

** Department of Physics, University of Leicester, Leicester LE1 7RH, UK

†† Rutherford Appleton Laboratory, Didcot, OX11 0QX, UK

We have discovered a new hard X-ray source near supernova 1987A, in the Large Magellanic Cloud, from the X-ray astronomy satellite Ginga. The present error box of $0.2^\circ \times 0.3^\circ$ includes the supernova. The energy spectrum is very hard above 10 keV and unusual for any of the known classes of X-ray source. The flux in the range 10–30 keV as of 2–3 September is $\sim 5 \times 10^{-11} \text{ erg cm}^{-2} \text{ s}^{-1}$. The source intensity increased steadily throughout July, August and early September, but an observation in late September revealed no further increase. The positional agreement, a steady brightening and an unusual hard spectrum support the identification of the source as SN1987A.

We have been regularly observing SN1987A from the X-ray astronomy satellite Ginga^{1,2} since 25 February 1987, immediately following the optical discovery^{2,3}. Ginga carries a set of proportional counters of 4,000 cm² effective area and a field of view of $2^\circ \times 4^\circ$ full width. SN1987A (hereafter abbreviated as SN) is $\sim 0.6^\circ$ away from LMC X-1, one of the brightest X-ray sources in the Large Magellanic Cloud (LMC) with an intensity of $6 \times 10^{-10} \text{ erg cm}^{-2} \text{ s}^{-1}$ in the 2–10 keV range. LMC X-1 is known to have a very soft spectrum compared to typical compact binary X-ray sources, but it has a hard X-ray tail extending over 10 keV and its intensity varies with time. To distinguish the SN from LMC X-1, we have used two different modes of observation: (1) slow scans along a path through the SN and LMC X-1, and (2) pointing observations at a position offset by $\sim 1^\circ$ from LMC X-1 on the side of the SN as shown in Fig. 1, which gives an exposure to the SN with half the maximum sensitivity and little contribution from LMC X-1.

Figure 2 shows an example of the histograms of the count rate with time in different energy bands, which were obtained from the scanning observations on 2 and 3 September. In the energy range below 10 keV, the counts from LMC X-1 dominate. On the other hand, as the energy increases the peak position of the count rate shifts towards the line position near the SN, clearly indicating the presence of a hard X-ray source⁴ separated from LMC X-1. We are certain that this source did not exist in

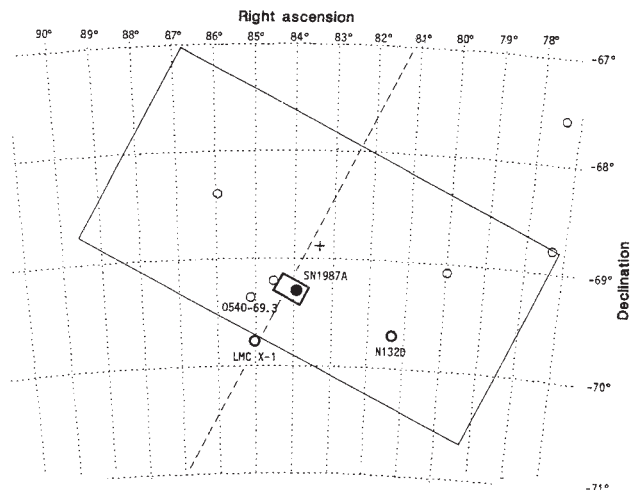


Fig. 1 Sky map of the LMC region. The scan path (dashed line) on 2–3 September, the target position (+) for the offset pointing mode and the full field of view ($2^\circ \times 4^\circ$) are indicated. The 90% confidence error box of the hard source (solid rectangle) is also shown.

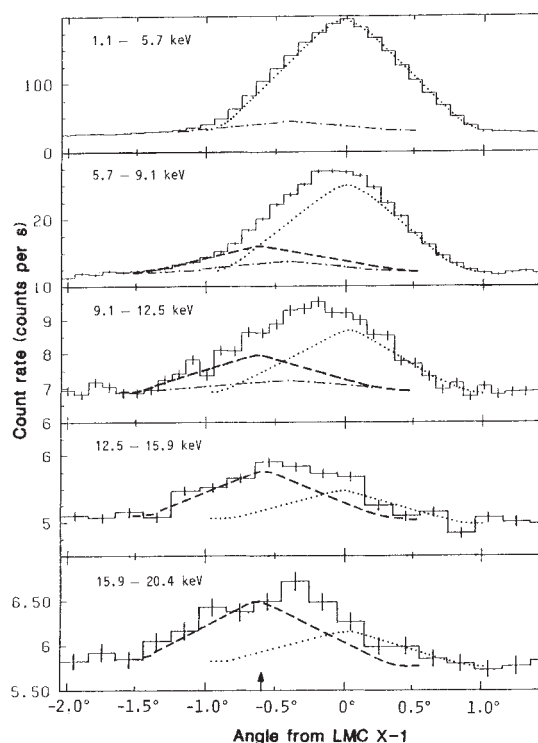


Fig. 2 Count-rate histograms obtained from the scans on 2–3 September are shown in five different energy bands. The dotted lines (LMC X-1), dash-dotted lines (SNR0540–69.3) and dashed lines (the hard source) indicate the fits to the collimator response function. The position angle of SN1987A is marked an arrow.

the March–April period (see Fig. 3). By fitting the collimator response function to the observed count-rate histograms, we determined separately the contribution of LMC X-1 and the intensity and line position of the hard X-ray source. The contribution of the nearby supernova remnant 0540–69.3 is also taken into account using the result of our observation of this remnant. The collimator response function was measured to an accuracy of 1 arc min before launch, and also verified in flight with the observation of the Crab nebula. The result of the fit is shown by the dashed lines in Fig. 2. Note that the contributions of LMC X-1 and SNR0540–69.3 are minor in the range above

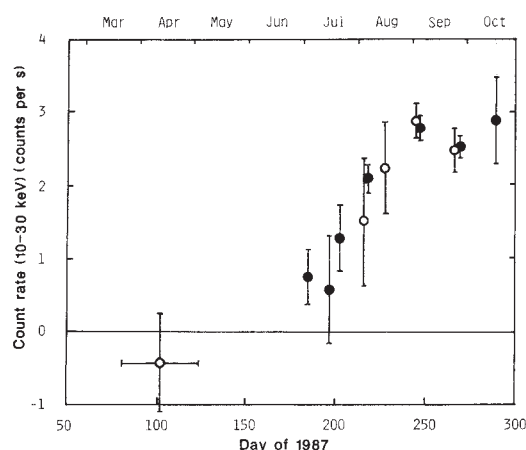


Fig. 3 The light curve of the hard source in terms of the count rate in the range 10–30 keV obtained from pointing (filled circle) and scanning (open circle) observations.

10 keV and quickly diminish with increasing energy, clearly showing that the new source has a very hard spectrum. The line position of the hard source thus determined is in agreement with that of the SN within $\pm 0.1^\circ$ (90% confidence limit).

In order to determine the error box of the position of the hard source, we conducted separate scans on 4–7 and 24–29 September along four different paths, parallel to the path of 2–3 September and separated from it by $\sim 0.5^\circ$ and 1° on both sides. From the comparison of the peak count rate observed on each scan path, we determined the 90% confidence error box of the hard source of size $0.2^\circ \times 0.3^\circ$, as shown in Fig. 1, which indicates the SN.

Figure 3 shows the light curve of the hard source in terms of the count rate (counts per cm^2 per s) in the range 10–30 keV. The result is consistent with a steady increase of the source intensity between June and early September. However, the observations of 24–26 September, three weeks after the preceding ones on 2–3 September, do not show a further intensity increase. At present, it is unclear whether or not the intensity is near a peak. The data appear inconsistent with a monotonic increase.

We find that the new source has a very unusual spectrum. Figure 4 shows the energy spectrum of the source obtained from a pointing observation on 3 September when the flux was highest, in the form of a photon number spectrum corrected for the energy-dependent detection efficiency. Two nearby supernova remnants, N132D and 0540–69.3, were in the field of view of these observations. The result of the LMC survey from the Einstein Observatory⁵ shows that the contribution from the sources in our field of view other than these two supernova remnants is negligible. Contributions of these two supernova remnants were subtracted using the results of separate observations of these supernova remnants. But the spectrum below 4 keV is still subject to uncertainty, because it is difficult to eliminate unambiguously a slight contamination from the brighter and variable source LMC X-1.

Figure 4 clearly shows that the shape of the spectrum is extraordinary. It becomes very hard at high energies and is essentially flat above 12 keV. This is unlike any of the known classes of X-ray sources. The observed spectrum can be explained by a composite of a thermal bremsstrahlung spectrum with energy kT (where k is the Boltzmann constant and T is temperature) of ~ 4 keV and a flat spectrum, independent of energy. The same amount of neutral absorbing matter was assumed for both components. There is no doubt that these two components come from the same source, because of the positional agreement and similar light curves observed for two energy ranges above and below 10 keV. We find little indication of a

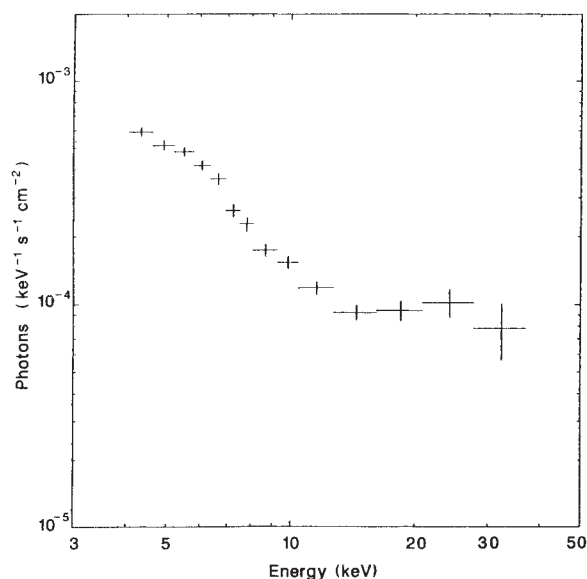


Fig. 4 Photon number spectrum of the hard source obtained from pointing observations on 3 September, corrected for the detection efficiency. The detection efficiency as a function of X-ray energy was calibrated before launch and after (using the Crab nebula). The iron line intensity for the spectral model used was $(1.3 \pm 0.3) \times 10^{-4}$ photons $\text{cm}^{-2} \text{s}^{-1}$.

low-energy cutoff due to photoelectric absorption down to 4 keV. This implies that the absorption column density of cool matter is not much larger than 10^{23} H atoms per cm^2 for cosmic abundances of elements. In addition, a signature of an iron emission line is present. The energy flux in the range 10–30 keV is $\sim 5 \times 10^{-11}$ erg $\text{cm}^{-2} \text{s}^{-1}$.

The entire LMC was previously surveyed from Einstein Observatory by Long, Helfand and Grabelsky⁵. The present error box still includes two faint, unidentified Einstein sources, but the positional agreement, a steady brightening and an unusual hard spectrum of the source argue for the hard source being SN1987A or directly related to the SN. If that is the case, the measured intensity as of 2–3 September corresponds to a luminosity of $\sim 1.5 \times 10^{37}$ erg s^{-1} in the range 10–30 keV for a distance of 50 kpc.

Note added in proof: The referee has highlighted an apparent discrepancy at 5 keV between our results and those in the accompanying paper by R. Sunyaev *et al.*⁶. We are certain that the soft X-ray component below 10 keV from the supernova was absent in the March–April period, when observed soft X-rays were entirely accounted for by the flux from N132D and SNR0540–69.3 in the same field of view. (We have separately measured the spectra of both sources.) The Ginga effective area as well as the detection efficiency in the energy range concerned are such that we can determine the intensity of the soft component unambiguously. It is possible that the soft component was significantly weaker during the measurements of Sunyaev *et al.*⁶.

Received 7 October; accepted 26 October 1987.

1. Makino, F. *et al. Astrophys. Lett.* **25**, 223 (1987).
2. Shelton, I. (reported by Kunkel, W. & Madore, B.) *IAU Circ. No.* 4316 (1987).
3. Jones, A. & Nelson (reported by Bateson, F. M.) *IAU Circ. No.* 4316 (1987).
4. Makino, F. *IAU Circ. No.* 4447 (1987).
5. Long, K. S., Helfand, D. J. & Grabelsky, D. A. *Astrophys. J.* **248**, 925 (1981).
6. Sunyaev, R. *et al. Nature* **330**, 227–229 (1987).

A search for soft X-rays from supernova 1987A

B. Aschenbach, U. G. Briel, E. Pfeffermann,
H. Bräuninger, H. Hippmann & J. Trümper

Max-Planck-Institut für Physik und Astrophysik,
Institut für extraterrestrische Physik, 8046 Garching, FRG

On 23 February 1987, supernova 1987A exploded in the Large Magellanic Cloud (LMC). This event offers the possibility of studying the evolution of a nearby supernova remnant from the earliest beginnings over the entire electromagnetic spectrum with unprecedented sensitivity. We searched for soft X-ray emission (0.2–2.1 keV) from SN1987A by means of a sounding rocket-borne experiment on 24 August. No positive signal was detected, and an upper limit for the luminosity of 1.5×10^{36} erg s⁻¹, between 0.2–2.1 keV at a 95% confidence level has been derived.

The scientific payload consisted of an imaging Wolter type I telescope with a position-sensitive proportional counter in the focal plane. The telescope and the X-ray detector were calibrated on the ground. The effective collecting area of the instrument peaks at 13 cm² for 1-keV X-rays with a spatial resolution of better than 1 arc min half-energy radius. The total field of view has a diameter of 45 arc min. Both the telescope and the detector were built ten years ago and further details of the design can be found in refs 1–3. The Skylark 7 rocket was launched on 24.69 August UT from Woomera, Australia, and reached a maximum altitude of 270 km. The viewing direction of the telescope was determined by two charge-coupled device-based star cameras, one almost co-aligned with and the other at right angles to the X-ray telescope. The telescope was first pointed at the known X-ray source LMC X-1, which is offset by ~40 arc min from the SN1987A position. LMC X-1 was clearly detected in real time at the expected flux level and close to the expected position. From this position a relative aspect solution accuracy of 1.2 arc min has been derived with an absolute pointing error of 1 arc min. The telescope was then slewed to the SN1987A position which it held for 240 s. Figure 1 shows the aspect-corrected X-ray image with the central cross indicating the position of SN1987A. From the absence of source counts an upper limit for the count rate of 0.16 mCrab (1.2×10^{-2} counts s⁻¹) has been derived at a confidence level of 95%. Assuming an absorbing column density of 6×10^{20} cm⁻² (ref. 4), the count rate upper limit corresponds to a flux of 10^{-12} erg cm⁻² s⁻¹ keV⁻¹ at 1 keV or a luminosity of 1.5×10^{36} erg s⁻¹ in the 0.2–2.1 keV band. These limits hold for a thermal spectrum with a temperature between 0.1–10 keV, or a non-thermal spectrum with photon number power spectral index between 0 and -3.

The X-ray production in a supernova event and in the very early phase of the associated supernova remnant may take place within the expanding envelope or outside it. Interior mechanisms include radiation from a central pulsar⁵ and Compton scattering of γ -rays from the radioactive decay of the nickel/cobalt core^{6,7}. But soft X-rays from these processes are likely to be significantly absorbed by the mass M_e inside the envelope. Assuming a maximum supernova velocity of $v_{ex} = 26,000$ km s⁻¹ (ref. 8) and uniform cold matter of cosmic abundances in the envelope, the optical depth against photoelectric absorption is $\tau = 5.4 \times M_e$ at 2 keV at the time of our observation, with M_e measured in solar masses.

The interaction of the expanding envelope with the circumstellar material can give rise to X-ray emission from the outer regions of the envelope. If the circumstellar environment is dominated by the stellar wind of the progenitor star, the thermal emission from the shocked supernova gas depends on the effective density, determined by the ratio A of the progenitor

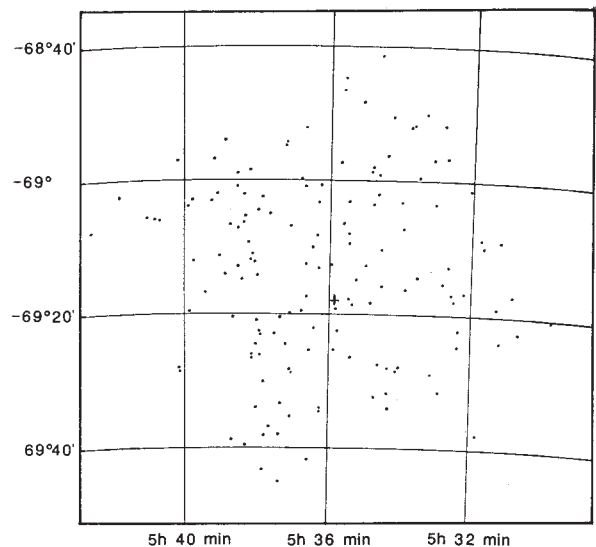


Fig. 1 Soft X-ray image of the SN1987A region at 24.69 August 1987 UT. The position of SN1987A is indicated by the light cross close to the image centre. Each dot represents a single X-ray photon or background event.

mass loss rate $\dot{M}$ to the wind velocity v_w . Following Chevalier and Fransson⁹, we have estimated the X-ray luminosity for $t_{ex} = 180$ days after the explosion to be $L_x = 6 \times 10^{34}$ A² erg s⁻¹ or 8.8×10^{35} A² erg s⁻¹ for a power-law radial density distribution of the supernova gas with an index of $m = 7$ or $m = 12$ respectively, where $A = (\dot{M}/v_w)/(8 \times 10^{-6} M_\odot \text{ yr}^{-1} \text{ per } 500 \text{ km s}^{-1})$. The X-ray temperature is between 3.5 and 7×10^7 K. From our observation we conclude that A is ≤ 2 or ≤ 7 , or $\dot{M} \leq 1.5 \times 10^{-5} M_\odot \text{ yr}^{-1}$ or $\dot{M} \leq 5.7 \times 10^{-5} M_\odot \text{ yr}^{-1}$ for $v_w = 500 \text{ km s}^{-1}$, for the cases $m = 12$ and $m = 7$ respectively. Using the continuity equation we calculate upper limits of the wind density n at the radial distance $R = v_{ex} \times t_{ex} = 4.1 \times 10^{16}$ cm to $n = 400 \text{ cm}^{-3}$ and $1,500 \text{ cm}^{-3}$ for the two values of m considered. This supports the view that the circumstellar density around SN1987A is low, and it is consistent with a low progenitor-mass-loss rate and a high wind velocity, as expected from a blue supergiant such as Sanduleak -69 202. Our upper limit clearly excludes circumstellar densities typical of winds from a red supergiant around SN1987A. If the progenitor star had passed through a red supergiant phase with an associated high mass-loss rate and low wind velocity v_{sw} , the supernova gas would have reached this region in a time $t = t_c = v_{sw}/v_{ex}$, if the wind had been terminated at a time t_c before the explosion¹⁰. For a red supergiant wind velocity of $v_{sw} = 10 \text{ km s}^{-1}$ we obtain a lower limit of $t_c = 1,300$ yr.

Inverse Compton scattering of optical photons off relativistic electrons can also produce X-ray emission. The number of relativistic electrons could be estimated from their synchrotron radio emission. But radio observations close to the time of our observation provide only upper limits of ~8 mJy at 2.3 GHz and 8.4 GHz (ref. 11), and therefore we cannot put significant constraints on either the relativistic electron population or the magnetic field.

Soft X-rays from a new supernova have been detected so far only from SN1980K in NGC6946 at a level of 3×10^{39} erg s⁻¹ about 35 days after maximum light¹². The ratio of the X-ray luminosity to the bolometric luminosity was $\sim 3 \times 10^{-4}$, whereas this ratio is $< 10^{-6}$ for SN1987A, making it a remarkably underluminous supernova in the soft X-ray region.

This first rocket observation of SN1987A was planned, prepared, organized, financed and carried out in the extraordinarily short timescale of five months. This was possible only with the effective cooperation of the Bundesministerium für Forschung und Technologie, the Australian DITAC, the DFVLR sections

in Cologne and Oberpfaffenhofen, and Dornier System, all of which are gratefully acknowledged.

Received 13 October; accepted 26 October 1987.

1. Trümper, J., Aschenbach, B. & Bräuninger, H. *SPIE Proc.* **184**, 12–19 (1979).
2. Aschenbach, B., Bräuninger, H., Ondrusch, A. & Trümper, J. *SPIE Proc.* **184**, 148–152 (1979).
3. Pfeffermann, E. & Briel, U. *Adv. Space Res.* **2**, 255–258 (1983).
4. Heiles, C. & Cleary, M. N. *Aust. J. Phys. Astrophys. Suppl.* **47**, 1–58 (1979).
5. Pacini, F. *ESO Conf. on SN1987A* Garching, July 1987.
6. McCray, R., Shull, J. M. & Sutherland, P. *Astrophys. J.* **317**, L73–L77 (1987).
7. Grebenev, S. A. & Sunyaev, R. A. *Soviet Astr. Lett.* **13**, 945–963 (1987).
8. Gry, C., Cassatella, A., Wamsteker, W. & Sanz, L. *IAU Circ. No.* 4324 (1987).
9. Chevalier, R. A. & Fransson, C. *Nature* **328**, 44–45 (1987).
10. Itoh, H., Hayakawa, S., Masai, K. & Nomoto, K. *Publ. astr. Soc. Japan* **39**, 529–537 (1987).
11. Norris, R. P. *et al.* *IAU Circ. No.* 4432 (1987).
12. Canizares, C. R., Kriss, G. A. & Feigelson, E. D. *Astrophys. J.* **253**, L17–L21 (1982).

X-rays expected from supernova 1987A compared with the source discovered by the Ginga satellite

M. Itoh*, S. Kumagai†, T. Shigeyama†, K. Nomoto† & J. Nishimura*

* Institute of Space and Astronautical Science, Tokyo 153, Japan

† Department of Earth Science and Astronomy, University of Tokyo, Tokyo 153, Japan

The X-ray astronomy satellite Ginga has detected X-rays from the region of supernova 1987A¹. The location and hard spectrum of this X-ray source make it a strong candidate for SN1987A^{1,2}. Among the possible mechanisms of X-ray emission from SN1987A^{3,4}, including the collision of the ejecta with circumstellar matter, non-thermal radiation from a pulsar, is Compton scattering of monoenergetic γ -rays from the radioactive decay $^{56}\text{Ni} \rightarrow ^{56}\text{Co} \rightarrow ^{56}\text{Fe}$ (refs 3, 5–7). It is clear that radioactive decays are powering the optical emission of SN1987A because its luminosity is now declining exponentially with the ^{56}Co decay rate^{8,9}. We have constructed hydrodynamic models of SN1987A and calculated the optical light curve and the X-ray light curve and spectrum from ^{56}Co decay. For the optical light curve, the model agrees better with the observations if some ^{56}Co is mixed into outer layers rather than being confined to the deepest layer of the ejecta. Such mixing has much larger effects on the X-ray light curve. Without the mixing of ^{56}Co , X-rays would not be observable by Ginga until ~ 250 d after the explosion. With mixing the X-rays emerge much earlier. X-ray emission from radioactive decay may therefore qualitatively account for the hard component of X-rays from SN1987A observed by Ginga, if ^{56}Co has been mixed into the outer layers. The soft component of the spectrum cannot be explained by the radioactive model and must be due to another mechanism.

The X-ray source detected by the Ginga satellite has several notable features. Its intensity increased after $t \sim 130$ d but seems to have levelled out during $t = 190$ – 220 d. (Here t is the time since the neutrino burst^{10,11}). Its spectrum is hard yet extends down to at least 4 keV with no indication of an energy cutoff². The spectrum could consist of separate soft and hard components. The soft component might be thermal radiation from the collision between the supernova ejecta and the circumstellar materials^{3,12,20}, and the hard component could be due to Compton degradation of γ -rays from ^{56}Co decay in SN1987A. This explanation for the hard component is examined here.

Several groups have calculated the expected X-ray flux and spectra resulting from radioactive decay^{5–7}. The line γ -rays are degraded into X-rays by Compton scattering. X-rays are then absorbed by heavy elements during the early phase of the expansion but later emerge from the surface layers when the expanding envelope becomes sufficiently thin. The time of X-ray

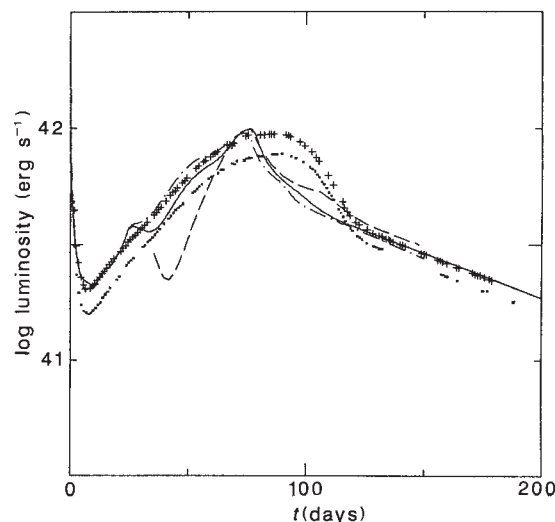


Fig. 1 Bolometric optical light curve for the supernova model with $E = 10^{51}$ erg and $M = 11.3 M_{\odot}$ ($M_{\text{env}} = 6.7 M_{\odot}$). For the dashed curve, ^{56}Co is assumed to be confined to the deepest layer, while the solid and dash-dotted curves assume that ^{56}Co is mixed uniformly up to the outer edge of helium layer ($M_r < 4.6 M_{\odot}$) and up to the middle of the hydrogen-rich envelope ($M_r < 6.0 M_{\odot}$) respectively. Production of $0.071 M_{\odot}$ ^{56}Ni is assumed. Observed points are taken from Catchpole *et al.* (+)⁸ and Hamuy *et al.* (x)⁹.

emergence is sensitive to the dynamical and chemical structure of the exploding star, depending on the expansion velocity, the mass of the ejecta, and abundances. The X-ray light curve and spectrum (and thus detectability with Ginga) are found to be highly model dependent, and thus construction of a good hydrodynamic model is crucial. Conversely, X-ray studies of SN1987A will provide a useful diagnostic tool of explosion models.

Our model of supernova ejecta consists of a heavy element core of $2.4 M_{\odot}$, where $M_{\odot}$ is the solar mass (composed of $0.07 M_{\odot}$ ^{56}Ni , $0.26 M_{\odot}$ Ar-S-Si and $2.07 M_{\odot}$ Mg-Ne-O-C), a $2.2 M_{\odot}$ helium-rich layer (4% with mass fraction in carbon and 1% in oxygen), and a hydrogen-rich envelope of mass M_{env} (including heavy elements at 0.25 times solar abundances)^{13,14}. Here we assume that the progenitor star was $\sim 20 M_{\odot}$ on its main-sequence, developed a helium core of $\sim 6 M_{\odot}$ (inferred from the known luminosity of Sk-69 202¹⁵), and formed a $1.4 M_{\odot}$ neutron star after the explosion. Then, the hydrodynamics of the supernova ejecta depends on M_{env} and the explosion energy E . These two are free parameters in our calculation because they are highly uncertain, depending respectively on the mass-loss history before the explosion and the mechanism of explosion. To choose the model parameters, we must resort to comparison of the hydrodynamic models with observations. Most useful is the comparison with the unique optical light curve.

For the optical light curve calculation, we applied a flux-limited diffusion approximation¹⁶. Figure 1 compares the bolometric light curve of SN1987A^{8,9} with several theoretical curves. Several attributes of the supernova can be deduced. The mass of ^{56}Ni produced by the explosion is $\sim 0.07 M_{\odot}$, the figure being derived from the luminosity of the exponential tail. The parameters of the best fit model shown in Fig. 1 are $E \approx 10^{51}$ erg and $M_{\text{env}} = 6.7 M_{\odot}$ (and thus $M = 11.3 M_{\odot}$). This model is used in the following X-ray studies. For smaller M_{env} ($= 2.4 M_{\odot}$), the light curve reaches a peak luminosity of $\sim 10^{42}$ erg s⁻¹ at $t \approx 50$ d, earlier than observed. The distribution of ^{56}Co is another important parameter for the optical light curve. During the shock propagation through the star, the expanding core materials are decelerated by the low density envelope. A reverse shock, and density inversion form. The matter in the core and the helium layer may undergo Rayleigh-Taylor instability^{14,15}, which could

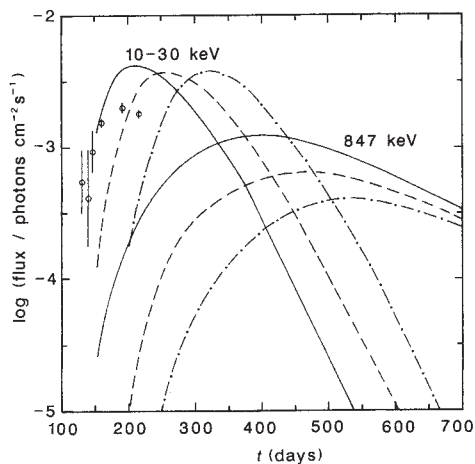


Fig. 2 Theoretical light curves of X rays in the 10–30 keV range and of the 847-keV γ -ray line for the three cases, OX (dash-dotted), where ^{56}Co is mixed up to the outer edge of the oxygen-rich core ($M_r = 2.4 M_\odot$), HE (dashed), where ^{56}Co is mixed up to the outer edge of helium layer ($M_r = 4.6 M_\odot$), and HY, where mixing takes place into part of the hydrogen-rich envelope ($M_r = 6.0 M_\odot$). A distance of 50 kpc is assumed. The open circles are the observed X-ray light curve in the 10–30 keV range.

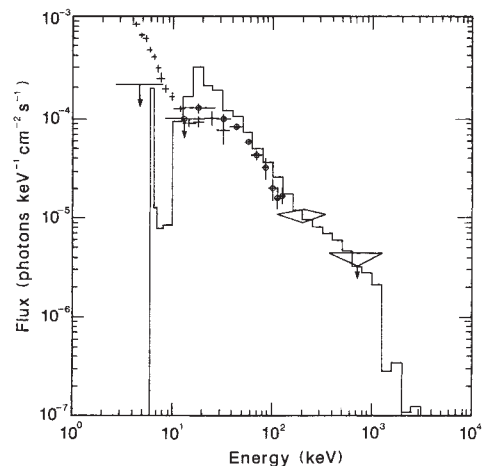


Fig. 3 The calculated spectrum at $t = 200$ d for case HY. The emission line around 6 keV is due to the $K\alpha$ fluorescence line of iron. The crosses indicate the spectrum observed by Ginga², and the open circles and the diamonds are obtained by Kvant²¹.

lead to mixing of ^{56}Ni – ^{56}Co into the outer layers. The extent of mixing is very uncertain and thus an unknown parameter in our calculation. The dashed, solid, and dash-dotted curves are the light curves for the case with no mixing, mixing up to $M_r = 4.6 M_\odot$, and up to $M_r = 6.0 M_\odot$ respectively. If mixing has taken place, the heat wave reaches the photosphere earlier so that the resulting light curve shows a smoother increase and a better fit to the observations (Fig. 1).

The optical light curve of SN1987A is equally well reproduced by the above model and Woosley's model 10H¹⁷. It should be noted that the optical depth of our model is smaller than that of Woosley's 10H at the same epoch because of our smaller M_{env} . For the X-ray light curve, such a difference may have important effects on the time of X-ray emergence and flux.

In obtaining X-ray and γ -ray light curves and spectra, we used a Monte Carlo method to simulate the propagation of photons in the expanding medium. Photons interact with matter primarily through Compton scattering and photoelectric absorption. Although our code can deal with electron-positron pair creation and coherent scattering by atoms, they are not important in this problem. For Compton scattering the Klein-Nishina differential cross-section was used. For almost-neutral atoms⁶, we used the cross-sections given in ref. 18. As most of the photons escape from the ejecta by much less than the expansion time, we assumed a stationary configuration of the medium. The velocity of atoms due to both thermal and bulk motion of the medium is also negligible in the scatterings. We took into account ^{56}Co -decay γ -ray lines that have emission probabilities $>10\%$.

We calculated X-ray light curves and spectra for several cases of mixing. Figure 2 shows the intensity changes in the 10–30 keV range and the 847 keV γ -ray line for three cases; case OX (dash-dotted), where ^{56}Co is uniformly mixed up the outer edge of the oxygen-rich core ($M_r = 2.4 M_\odot$); case HE (dashed), where ^{56}Co is mixed up the other edge of the helium layer ($M_r = 4.6 M_\odot$); and case HY, where mixing takes place into part of the hydrogen-rich envelope ($M_r = 6.0 M_\odot$). The Thomson scattering optical depths at the outer edge of the layer that includes cobalt are $\tau_{\text{Th}} = 5.3$ (case OX), 2.4 (case HE) and 1.3 (case HY) at $t = 200$ d. They are significantly smaller than $\tau_{\text{Th}} = 8.7$ when ^{56}Co is confined to the central layer at $M_r \approx 0.07 M_\odot$. The above optical depth is scaled as $(200 \text{ d}/t)^2$.

Two points are apparent from Fig. 2. (1) X- and γ -rays emerge earlier with more extensive mixing because of smaller τ_{Th} in the cobalt layer (see also ref. 19). Accordingly, the intensity peak of the X-ray comes in the order $t = 200$ d, 250 d and 320 d for cases HY, HE and OX respectively. (2) The peak X-ray flux does not depend much on the degree of mixing. If ^{56}Co alone is mixed, the peak flux is larger for earlier peaks because of the larger amount of cobalt, but with uniform mixing of all the core material into the helium- and hydrogen-rich layers, larger photoelectric absorption decreases the flux, which almost cancels the effect of the larger amount of cobalt.

Compare these light curves with the Ginga observations². The X-ray emergence observed around $t \sim 130$ d is consistent with case HY, while the observed flux at $t = 190$ – 220 d is close to that of case HE. Case OX is clearly inconsistent with the observation. This implies that mixing of ^{56}Co up to $M_r = 4.6$ – $6.0 M_\odot$ is required to account for the Ginga X-ray observation with the adopted hydrodynamic model. If we assume that mixing is limited only through the helium layer as in case HE because of the low density and high velocity of the hydrogen-rich envelope, then models with slightly smaller M_{env} and smaller E than our model might be in better agreement with observation.

Figure 3 shows the calculated spectrum at $t = 200$ d for case HY. The observed spectrum below 12 keV (ref. 2) cannot be accounted for by the radioactive model because of the photoelectric absorption of low energy X-rays. The spectrum above 12 keV is generally in good agreement with observations by Ginga² and Kvant²¹. In particular, the power law spectrum, $E^{-1.4}$, at 30–200 keV shows an excellent fit to the observed spectrum²¹. On the other hand, the peak flux at ~ 20 keV for case HY is about a factor of 2–3 too high. For case HE, the calculated spectrum is in good accord with the observed one at 10–30 keV, while the high-energy portion of the spectrum is a little too low. Again mixing of ^{56}Co up to $M_r = 4.6$ – $6.0 M_\odot$ is necessary to explain the observed spectrum.

One may argue that models without mixing could be consistent with the early detection of X-rays by Ginga if M_{env} is much smaller or E is much larger. In fact, for $M_{\text{env}} = 2.4 M_\odot$ ($M = 7.0 M_\odot$), early observation is possible without mixing of ^{56}Co . Also, we found empirically and derived analytically the result that, for the same M_{env} , the time of the X-ray intensity peak t_{max} scales as $E^{-1/2}$. For these cases, however, the intensity increases too steeply, the peak flux is unacceptably too high, and the optical light curve reaches its peak too early. Models with mixing thus seem more likely.

Although there remain some quantitative discrepancies

between our simple models and the observations, the theoretical flux is subject to large uncertainties. In particular, the mixing process would involve unknown effects due to, for example, the changes in density and abundance distribution, formation of the clumpy medium. Further observations will provide us with some information about these effects.

Finally, if the mixing of ^{56}Co up to $M_r = 4.6\text{--}6.0 M_\odot$ has actually taken place, the prediction is that the 847-keV γ -ray line flux will reach 4×10^{-4} photons $\text{cm}^{-2} \text{s}^{-1}$ around $t = 240\text{--}330$ d and thus will be detected by the planned balloon observations. Detection of hard X-rays is also expected around that time.

We would like to thank K. Koyama, F. Makino, Y. Tanaka, H. Inoue and the Ginga team for the exciting information of their observations. We are grateful to Drs T. Ebisuzaki and N. Shibazaki for stimulating correspondence. This work has been supported in part by the Grant-in-Aid for Scientific Research of the Ministry of Education, Science and Culture in Japan and by the Space Data Analysis Center, Institute of Space and Astronautical Sciences.

Note added in proof: After submitting this paper, we received preprints on the Kvant observation²¹ and on the theoretical work which also examined the effect of mixing and reached essentially

the same conclusions as ours²². In Fig. 3, we plotted the observed points by Kvant²¹ for comparison with our models.

Received 9 October; accepted 23 October 1987.

1. Makino, F. *IAU Circ. No. 4447* (1987).
2. Dotani, T. *et al. Nature* **330**, 230–232 (1987).
3. McCray, R., Shull, J. M. & Sutherland, P. *Astrophys. J.* **317**, L73–L77 (1987).
4. Itoh, H., Hayakawa, S., Masai, K. & Nomoto, K. *Publ. astr. Soc. Japan* **39**, 529–537 (1987).
5. Gehrels, N., MacCallum, C. J. & Leventhal, M. *Astrophys. J.* **320**, L19–L22 (1987).
6. Xu, Y., Sutherland, F. & McCray, R. *Astrophys. J.* (in the press).
7. Ebisuzaki, T. & Shibazaki, N. *Astrophys. J.* (submitted).
8. Catchpole, R. *et al. Mon. Not. R. astr. Soc.* (in the press).
9. Hamuy, M., Suntzeff, N. B., Gonzalez, R. & Martin, G. *Astr. J.* (in the press).
10. Hirata, K. *et al. Phys. Rev. Lett.* **58**, 1490–1493 (1987).
11. Bionta, R. M. *et al. Phys. Rev. Lett.* **58**, 1094–1096 (1987).
12. Nomoto, K., Shigeyama, T., Hayakawa, S., Itoh, H. & Masai, K. in *Proc. 4th George Mason Workshop: Supernova 1987A in the LMC* (ed. Kafatos, M.) (Cambridge University Press in the press).
13. Shigeyama, T., Nomoto, K., Hashimoto, M. & Sugimoto, D. *Nature* **328**, 320–323 (1987).
14. Nomoto, K., Shigeyama, T. & Hashimoto, M. in *Proc. ESO Workshop on SN1987A* (ed. Danziger, J.) (ESO, Garching, 1987).
15. Woosley, S. E., Pinto, P. & Ensmann, L. *Astrophys. J.* (in the press).
16. Pomraning, G. C. *Astrophys. J.* **266**, 841–847 (1983).
17. Woosley, S. E. *Astrophys. J.* (submitted).
18. McMaster, W. H. *et al. Compilation of X-ray Cross-sections*, Sect. II, Rev. 1 (Livermore, California, 1969).
19. Ebisuzaki, T. & Shibazaki, N. *Astrophys. J.* (submitted).
20. Masai, K., Hayakawa, S., Itoh, H. & Nomoto, K. *Nature* **330**, 235–236 (1987).
21. Sunyaev, R. *et al. Nature* **330**, 227–229 (1987).
22. Pinto, P. & Woosley, S. E. *Astrophys. J.* (submitted).

Thermal X-ray emission from supernova 1987A

K. Masai*, S. Hayakawa†, H. Itoh‡ & K. Nomoto§

* Institute of Plasma Physics, Nagoya University, Nagoya 464, Japan

† Department of Astrophysics, Nagoya University, Nagoya 464, Japan

‡ Department of Astronomy, University of Kyoto, Kyoto 606, Japan

§ Department of Earth Science and Astronomy, University of Tokyo, Tokyo 153, Japan

A new hard-X-ray source has recently been discovered in the region of SN1987A by the X-ray astronomy satellite Ginga¹. Several mechanisms are possible for X-ray emission from SN1987A. Non-thermal X-rays are expected from the Compton degradation of γ -rays from ^{56}Co decays^{2–4} and the radiation of a pulsar left inside^{2,5}. On the other hand, thermal X-rays are expected from the interaction of the expanding ejecta with circumstellar matter^{6–10}. Here we show that the observed X-ray spectrum can be accounted for by a superposition of thermal X-rays and Compton degradation of γ -rays, and discuss some constraints on physical quantities related to the mass-loss history of the progenitor.

The observed spectrum is characterized by a component below 10 keV which decreases with increasing photon energy, and a component above 10 keV which is nearly flat. This unusual X-ray spectrum may be understood as follows: X-rays below 10 keV are likely to be due to thermal emission from the shock-heated ejecta, and those above 10 keV to γ -ray degradation inside the ejecta. If thermal emission due to the collision of the ejecta with circumstellar matter is responsible for X-rays below 10 keV, the epoch of the collision can be estimated to be ~ 0.2 yr after the explosion if ≈ 0.5 yr is the time when the X-ray flux at ~ 10 keV reaches its maximum. The observed X-ray light curve then requires the inner radius of circumstellar matter to be $\sim 1 \times 10^{16}$ cm for an expansion velocity $V_{\text{ex}} \approx 2 \times 10^9 \text{ cm s}^{-1}$. This model may be confirmed by observations of infrared emission from shock-heated dust¹².

Drawing on previous work^{7–11}, we sketch the mechanism of thermal X-ray emission due to interaction between the ejecta and circumstellar matter. X-rays emitted from the reverse-shocked ejecta dominate those from the blast-shocked circumstellar matter because of a much higher density in the ejecta.

The electron temperature of the shocked ejecta is raised to > 10 keV and then decreases with time to ~ 10 keV at $\sim t_{\text{max}}$, at which the photon flux at 10 keV reaches its maximum. Free-free emission dominates the thermal spectrum except for a contribution of iron K-line emission around 7 keV.

If the circumstellar matter concerned is the remnant of a stellar wind, the thermal X-ray luminosity can be estimated from the mass loss rate $\dot{M}_w$, the wind velocity v_w and the time τ since the progenitor left the mass-loss phase. Thus, the photon flux at ~ 10 keV is expected near t_{max} to be

$$I(10 \text{ keV}) \approx 3 \times 10^{-4} (\dot{M}_w / 10^{-5} M_\odot \text{ yr}^{-1})^2 (v_w / 10^6 \text{ cm s}^{-1})^{-3} \times (\tau / 10^3 \text{ yr})^{-(1+\beta)} \text{ photons cm}^{-2} \text{ keV}^{-1} \text{ s}^{-1} \quad (1)$$

where the distance to SN1987A is assumed to be 55 kpc, and β is a numerical factor depending on the ejecta model. We assume $\beta = 1/2$ hereafter, according to our previous numerical calculations. From the radius of the ejecta at the moment of the collision with the wind remnant we derive a relation

$$(v_w / 10^6 \text{ cm s}^{-1}) (\tau / 10^3 \text{ yr}) \approx 0.4 (V_{\text{ex}} / 2 \times 10^9 \text{ cm s}^{-1}) (t_{\text{max}} / 0.5 \text{ yr}) \quad (2)$$

where the quantities on the right hand side can be determined observationally. Radio observations showed that the emission at 1.4 GHz reached a maximum a few days after the explosion¹³. The early detection of radio emission gives a constraint on the density in the stellar wind remnant. For the free-free opacity at a frequency ν to be less than unity, we have

$$(\dot{M}_w / 10^{-5} M_\odot \text{ yr}^{-1}) (v_w / 10^6 \text{ cm s}^{-1})^{-1} < 0.3 (V_{\text{ex}} / 2 \times 10^9 \text{ cm s}^{-1})^{3/2} (t_{\text{max}} / 0.5 \text{ yr})^{3/2} (\nu / 1 \text{ GHz}) (T_e / 10^4 \text{ K})^{3/4} \quad (3)$$

for fully ionized gas of electron temperature T_e in the wind remnant. Thus, from the observed thermal flux at 10 keV, we can estimate the mass-loss rate as

$$\dot{M}_w \approx 3 \times 10^{-6} (v_w / 10^6 \text{ cm s}^{-1})^{3/4} \times (I(10 \text{ keV}) / 10^{-4} \text{ photons cm}^{-2} \text{ keV}^{-1} \text{ s}^{-1})^{1/2} \times (V_{\text{ex}} / 2 \times 10^9 \text{ cm s}^{-1})^{3/4} (t_{\text{max}} / 0.5 \text{ yr})^{3/4} M_\odot \text{ yr}^{-1} \quad (4)$$

We performed a numerical calculation of the dynamical evolution and non-equilibrium X-ray emission for the

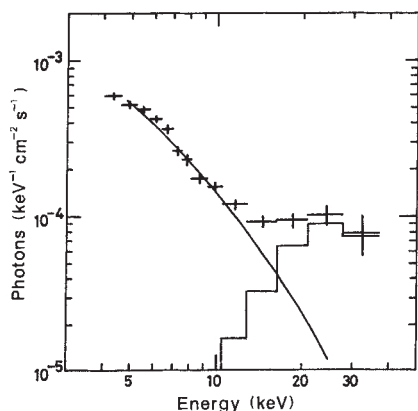


Fig. 1 Calculated thermal spectrum (thick solid line) with an energy resolution of $0.49 (E/\text{keV})^{1/2}$ keV (FWHM) is over-laid on the observed spectrum with Ginga¹ (crosses). The residual flux over the thermal emission is represented by the histogram. Contributions of free-free, free-bound, 2-photon decay and bound-bound processes are represented by thin solid lines. $Z = 1/3Z_{\odot}$, the hydrogen column density of $4 \times 10^{21} \text{ cm}^{-2}$ and the distance of 55 kpc to SN1987A are assumed.

collision at $1.1 \times 10^{16} \text{ cm}$, which is the inner radius of the circumstellar matter. The density of the circumstellar matter was taken to be $\approx 6.9 \times 10^4 (r/1.1 \times 10^{16} \text{ cm})^{-2} \text{ cm}^{-3}$ at a distance r from SN1987A, corresponding to $\dot{M}_w/v_w \approx 2.8 \times 10^{-6} M_{\odot} \text{ yr}^{-1}/10^6 \text{ cm s}^{-1}$. The calculated X-ray spectrum at $t_{\text{max}} \sim 0.5 \text{ yr}$ is compared with the spectrum observed by Ginga in Fig. 1. The observed spectrum below 10 keV is reproduced by thermal emission from shock-heated materials. However, the photon flux at 1 keV expected from the above model is higher than the upper limit obtained by a rocket experiment¹⁴.

For a blue supergiant with a typical wind velocity ≈ 100 – $1,000 \text{ km s}^{-1}$, the value of $\dot{M}_w$ given by equation (4) is too large. On the other hand, for $v_w \approx 30 \text{ km s}^{-1}$, the mass-loss rate is estimated to be $\dot{M}_w \approx 7 \times 10^{-6} M_{\odot} \text{ yr}^{-1}$, which is not inconsistent with what is expected for a red supergiant. This result implies that the circumstellar matter may be attributed to the stellar wind in the latest phase of a red supergiant or in the transition phase to a blue supergiant. Then the time which has passed since the stellar wind ceased is estimated from equation (2) as $\tau \approx 100 (v_w/3 \times 10^6 \text{ cm s}^{-1})^{-1} \text{ yr}$. In any case, the lifetime in the blue supergiant phase would be of the order of 100 yr if X-rays below 10 keV arise from the interaction between the ejecta and the wind remnant.

The residual flux above the thermal spectrum is shown in Fig. 1. Because of strong absorption below 20 keV, this residual spectrum may be ascribed to X-rays coming from the optically thick ejecta through Compton degradation of γ -rays. The cutoff energy at which the optical depth for photoabsorption becomes unity is approximately².

$$E_c \approx 30(Z/0.2 Z_{\odot})^{1/3} (M_e/6 M_{\odot})^{2/3} (V_{\text{ex}}/8 \times 10^8 \text{ cm s}^{-1})^{-4/3} \times (t/0.5 \text{ yr})^{-4/3} \text{ keV} \quad (5)$$

Here Z is the metallicity of the envelope of the ejecta and $Z_{\odot}$ is the solar value, M_e is the envelope mass, and t is the time since the explosion. X-rays from Compton degradation reach a maximum for the scattering optical depth of about 6 (ref. 2) which happens at time

$$t \approx 0.5(M_e/6 M_{\odot})^{1/2} (V_{\text{ex}}/8 \times 10^8 \text{ cm s}^{-1})^{-1} \text{ yr} \quad (6)$$

The emergent X-ray flux due to Compton degradation of γ -rays depends strongly on the density structure and the chemical composition⁴. A realistic ejecta model is needed to account for the light curve and the spectrum in detail. Our simple estimate nevertheless suggests that the component above 10 keV is essen-

tially in agreement with the expected spectrum from Compton degradation.

In conclusion, the observed spectrum consists of thermal and Compton-degraded components. The thermal X-rays can be attributed to the shocked ejecta colliding with some circumstellar matter. If the circumstellar matter is a stellar wind remnant, the early detection of thermal X-rays may indicate a short lifetime of the progenitor in its blue supergiant phase. The light curve of thermal X-rays depends on the density structure of the circumstellar matter and can provide information on it. Continual watching by active instruments^{1,15} is desired.

Received 19 October; accepted 6 November 1987.

1. Dotani, T. *et al. Nature* **330**, 230–232 (1987).
2. McCray, R., Shull, J. M. & Sutherland, P. *Astrophys. J.* **317**, L73–L (1987).
3. Gehrels, N., MacCallum, C. J. & Leventhal, M. *Astrophys. J.* **320**, L19–L22 (1987).
4. Itoh, M., Kumagi, S., Shigeyama, T., Nomoto, K. & Nishimura, J. *Nature* **330**, 233–234 (1987).
5. Ostriker, J. P. *Nature* **327**, 287–288 (1987).
6. Chevalier, R. A. & Fransson, C. *Nature* **328**, 44–45 (1987).
7. Itoh, H., Masai, K. & Nomoto, K. in *IAU Colloquium 101, The Interaction of Supernova Remnants with The Interstellar Medium* (in the press).
8. Itoh, H., Hayakawa, S., Masai, K. & Nomoto, K. *Publs astr. Soc. Japan* **39**, 529–537 (1987).
9. Nomoto, K., Hayakawa, S., Itoh, H., Masai, K. & Shigeyama, T. in *Proc. ESO workshop on 1987A* (ed. Danziger, J.) (ESO, Garching, 1987).
10. Masai, K., Hayakawa, S., Itoh, H., Nomoto, K. & Shigeyama, R. in *Proc. IAU Colloquium 108, Atmospheric Diagnostics of Stellar Evolution* (in the press).
11. Nomoto, K., Shigeyama, T., Hayakawa, S., Itoh, H. & Masai, K. in *Proc. 4th George Mason Workshop SN1987A* (in the press).
12. Itoh, H. *Publs astr. Soc. Japan* (submitted).
13. Tuttle, A. J. *et al. Nature* **327**, 38–40 (1987).
14. Briel, U. G., Peffermann, E., Aschenbach, B., Bräuninger, H. & Trümper, J. *IAU Circ. No.* 4452 (1987).
15. Sunyaev, R. *et al. Nature* **330**, 227–230 (1987).

Neutron tori and the origin of r-process elements

C. J. Hogan* & J. H. Applegate†

* Steward Observatory, University of Arizona, Tucson, Arizona 85721, USA

† Department of Astronomy, Columbia University, New York, New York 10027, USA

If an accretion disk gets hotter than a few MeV, nuclei in the infalling matter are dissociated into their constituent neutrons and protons. Neutrons released by dissociation of matter falling at high accretion rates into a black hole or neutron star accumulate in a dense 'neutron torus'. Matter in accretion disks around compact objects like those responsible for galactic X-ray sources may thus provide ideal conditions for classical rapid or r-process nucleosynthesis¹. Systems in which accretion powers an outflow from a region near the compact object might thereby enrich the interstellar medium in r-process elements. These elements are usually thought to originate in other stellar environments, such as helium shell-flash, supernovae and tidally dispersed neutron stars^{2–5}. Nucleosynthesis of lighter elements (C, N, O...) in accretion disks has been discussed recently in refs 6 and 7; the r-process mechanism considered here was inspired by some ideas^{8,9} for manufacturing deuterium in pregalactic accretion disks.

Suppose that an accretion disk around a compact object of mass M reaches high enough temperatures ($\sim 10 \text{ MeV}$) that infalling nuclei (particularly helium, because of its high abundance) are dissociated into protons and neutrons. (Evidence has appeared in γ -ray lines that free neutrons are indeed liberated around compact objects⁹.) Both species will have energies of several MeV, but the two components are effectively dynamically decoupled. As the neutrons are not electrically charged they do not couple directly to the magnetic field, but the field renders the protons' motion fluid-like, and contributes the shear stresses which allows them to lose angular momentum and fall onto the compact object¹⁰. Instead, neutrons experience their own

dynamic viscosity which causes some to accrete and others to escape from the system. For large accretion rates, neutrons cannot escape as rapidly as they are supplied, so they accumulate in a torus where they remain in orbit for a decay time $\sim 1,000$ s. This is much longer than the infall time of protons so a dense reservoir of neutrons builds up. If 'seed' nuclei (iron peak elements) encounter this reservoir before being ejected in a hot coronal wind or jet, the system can act as a source of r-process elements. Although the ejection mechanism is very uncertain (and is not discussed here), many accreting systems are in fact observed to produce some sort of outflow. Under certain conditions the neutrons themselves might drive a wind. It is also possible that for some range of $\dot{M}$ lowering of the adiabatic index by nuclear dissociation may destabilize the disk and produce instabilities, in the same way that hydrogen ionization is responsible for dwarf nova outbursts¹⁶. The main uncertainty in disks acting as r-process sites is whether the material destined for outflow spends enough time in the neutron torus before it is ejected.

Let r_s denote the Schwarzschild radius of the accreting object. The infall time for protons is then¹⁰

$$t_{\text{infall}} \approx (r_s/c) \alpha^{-1} (r/r_s)^{3/2} (h/r)^{-2} \approx 10^{-5.5} \alpha^{-1} (r/r_s)^{3/2} (h/r)^{-2} (M/M_\odot) \text{ s}, \quad (1)$$

where α is the usual stress parameter (viscous stress divided by pressure) and $h/r \approx (V_{\text{thermal}}/V_{\text{escape}}) < 1$ is the local aspect ratio of the disk. Let $\dot{m}$ denote the accretion rate in units of the critical Eddington rate $\dot{M}_{\text{Ed}} = L_{\text{Ed}}/c^2 = 2\pi r_s m_p c/\sigma_T$. The density of protons is given by

$$n_p \approx (3/4\pi) \dot{M} t_{\text{infall}} m_p^{-1} h^{-1} r^{-2} \approx \dot{m} \left(\frac{M}{M_\odot} \right)^{-1} \left(\frac{h}{r} \right)^{-3} \left(\frac{r}{r_s} \right)^{-3/2} \alpha^{-1} \times 10^{20} \text{ cm}^{-3} \quad (2)$$

In a self-consistent model, the parameters $\dot{m}$, α , h/r and temperature T , are not independent, but models for disk accretion flows in galactic X-ray binaries which include realistic cooling processes and produce a good fit to the observed X-ray spectrum are often found to produce ion temperatures well above 10 MeV, sufficient to dissociate infalling nuclei^{11,12}.

In thermodynamic equilibrium, nuclei would begin to fall apart at $T \approx 300$ keV, in spite of the ~ 10 MeV binding of neutrons. But the radiation field in ion pressure-supported disks is far below equilibrium intensity, and dissociation occurs only when the probability of dislodging a neutron becomes appreciable in a single ion-ion collision, at ~ 10 MeV. The column density of proton matter intersected per orbit is $\sim hn_p$. If we evaluate this quantity at the radius r_n , somewhere near the radius $r_{10} = 10^2 r_s (h/r)^2$ at which $kT \approx 10$ MeV, where the neutrons are actually produced, we obtain

$$hn_p \approx \dot{m} \left(\frac{h}{r} \right)^{-3} \left(\frac{r_n}{r_{10}} \right)^{-1/2} \alpha^{-1} \times 10^{24} \text{ cm}^{-2} \quad (3)$$

With a cross-section $\sigma_{np} \approx 1$ barn, a neutron is likely to scatter $\sim 10^{-24} hn_p$ times per orbit. These collisions will cause the neutrons to spread towards both the inner and outer parts of the disk, and to populate the corona. The collisions are not sufficiently rapid however to advect neutrons into the accretion flow onto the compact object.

Once neutrons are liberated they form a disk or torus which evolves quite separately from the proton/electron disk. As shown below, the neutron torus can be much denser than the regular disk. Neutron collisions in the torus occur many times per orbit, contributing a classical dynamic viscosity which is independent of density, and hence producing an infall rate in the torus also independent of density:

$$\dot{M}_n \approx m_n r V_{\text{thermal}} \sigma_{nn}^{-1} (h/r)_n, \quad (4)$$

where $(h/r)_n$ is the aspect ratio for the neutron torus. As the

neutron infall is practically dissipationless, this $\dot{M}_n$ actually corresponds to a local spreading rate for the neutron torus. Infall can occur because some neutrons are evaporated. The mass and hence the density of neutrons in the torus is determined by $\dot{M}_{\text{decay}} + \dot{M}_n = \dot{M}_{\text{supply}}$, where the latter is the rate at which the infalling matter produces neutrons. As $\sim 10\%$ of infalling matter is neutrons, $\dot{M}_{\text{supply}} \approx 0.1 \dot{M}$. If the supply of neutrons $\dot{M}_{\text{supply}}$ exceeds $\dot{M}_n$, then the neutrons accumulate for a decay time; if $\dot{M}_{\text{supply}} < \dot{M}_n$, then the neutrons disappear by spreading as fast as they are produced. The difference in neutron density between the two cases is an enormous factor, $t_{\text{infall}}/t_{\text{decay}}$. Defining $m_n \equiv \dot{M}_n/\dot{M}_{\text{Ed}}$, we find

$$\dot{m}_n \approx \frac{1}{2\pi} \left(\frac{r}{r_s} \right)^{1/2} \left(\frac{\sigma_T}{\sigma_{nn}} \right) \left(\frac{h}{r} \right)^2 \quad (5)$$

As long as $\dot{M}_n(r)$ exceeds $\dot{M}_{\text{supply}}$ the neutron supply can be accommodated and the neutrons do not build up. As r decreases, h/r must increase for the $\dot{M}_n(r) = \dot{M}_{\text{supply}}$ solution to hold. If $\dot{M}_{\text{supply}}$ is large enough then even with $h \approx r$, $\dot{M}_n < \dot{M}_{\text{supply}}$ and the neutrons begin to accumulate, continuing until $\dot{M}_{\text{decay}} = m_n N_n/\tau_n = \dot{M}_{\text{supply}}$ is reached. Note that $\dot{M}_n$ is independent of neutron density so that $\dot{M}_n$ remains constant as the neutrons accumulate.

Disks with large accretion rates $\dot{m} \approx 10$ are therefore likely to grow a dense neutron torus, while those with $\dot{m} \ll 1$ are not. In general, $\dot{M}_n$ is not negligible compared to $\dot{M}_{\text{supply}}$, so an appreciable fraction of neutrons spread to larger or smaller radii before decaying. This is important for nucleosynthesis because those neutrons of 3–10 MeV spreading to larger radii will soon cool adiabatically to temperatures $\lesssim 1$ MeV, where neutron absorption cross-sections become larger (see below).

In the dense-torus case it is possible that 'neutron viscosity' dominates the α of the ion-electron disk as well. As $\dot{m}_n$ is at a minimum near r_s , one expects neutrons in the torus to pile up at a small radius. Some neutrons at $r \gg r_s$ will always form an evaporating wind. At large radii, the density decreases and eventually collisions become sufficiently rare that neutrons simply follow stellar-dynamical trajectories, bound or unbound, until they decay.

When $\dot{M}_{\text{supply}} > \dot{M}_n$, the mass of the torus grows until $\dot{M}_{\text{supply}} \approx \dot{M}_{\text{decay}}$ and each neutron lasts in the torus for a decay time. Then the neutron density is approximately

$$n_n \approx 0.1 n_p (1,000 \text{ s}/t_{\text{infall}}) \quad (6)$$

and the neutron density at r_n is given by

$$n_n \approx \dot{m} \left(\frac{M}{M_\odot} \right)^{-2} \left(\frac{h}{r} \right)^{-1} \left(\frac{r_n}{r_s} \right)^{-3} \times 10^{27.5} \text{ cm}^{-3} \quad (7a)$$

$$\approx \dot{m} \left(\frac{M}{M_\odot} \right)^{-2} \left(\frac{h}{r} \right)^{-7} \left(\frac{r_n}{r_{10}} \right)^{-3} \times 10^{21.5} \text{ cm}^{-3} \quad (7b)$$

Adiabatically cooled neutrons will be at a lower density $\propto T^{3/2}$. Notice that the uncertain parameter α has cancelled here, although there remains a very strong dependence on h/r . Equation (7) assumes that h/r of the neutron disk is the same as that of the proton disk, but because the neutron disk is likely to puff up from collisions, n_n is smaller by a factor (h/r) than given in these expressions. Even for a fat proton disk with $h/r \approx 1$, however, the neutron density is high enough to produce r-process elements.

Two criteria must be satisfied to produce r-process elements: sufficiently rapid neutron addition, and a sufficient total neutron exposure. A seed nucleus absorbs neutrons at a rate $n\sigma V_{\text{thermal}} \approx 10^{3.5} (n_n/10^{20} \text{ cm}^{-3}) \text{ s}^{-1}$, fairly independently of V_{thermal} as long as $kT \lesssim 1$ MeV. Adding ~ 200 neutrons to build the heaviest r-process elements takes only $\tau_{\text{nadd}} \approx 0.1 \text{ s}$ ($n_n/10^{20} \text{ cm}^{-3}$). The 'rapidity' requirement for a classical r-process is that $\tau_{\text{nadd}} < \tau_\beta$, where τ_β is the sum of the beta-decay half-lives of elements along the r-process path. Estimates give $0.1 \text{ s} \leq \tau_\beta \leq 10 \text{ s}$, so this

requirement appears to be easily satisfied in our neutron tori (although disks around supermassive black holes would not have sufficiently high densities). The second requirement is that a seed nucleus actually be exposed to neutrons for a long enough time, $\geq \tau_\beta$, for build-up to occur. A natural timescale to take is the proton infall time (1),

$$t_{\text{infall}} \approx 0.3 \text{ s} (\alpha/10^{-2})^{-1} (r/r_{10})^{3/2} (h/r) M/M_\odot \quad (8)$$

which exceeds τ_β for $r > r_{10}$. As the r-process is likely to occur for $r > r_{10}$, there is probably no fundamental difficulty in meeting the exposure requirement. This timescale may not be appropriate, however, as the interesting r-process ought to be occurring in outflowing, rather than infalling, matter. The exposure requirement is that matter ultimately destined for an outflow must reside in the cool part (≤ 1 MeV) of the neutron torus for at least a time τ_β . If this occurs then such outflows from accretion disks may enrich the galaxy in r-process elements.

If we assume that $10^{11} M_\odot$ of material in the disk of our galaxy is enriched to a mass fraction of 10^{-7} in r-process elements (a number fraction $\sim 10^{-11}$ in each of 50 mass-200 elements, say), then $10^4 M_\odot$ of r-process elements must be produced in the 10^{10} -year history of the galaxy. Each disk produces a mass ϵM of such elements, where the efficiency $\epsilon < 10\%$ (and possibly much less), and lasts for a time $\sim 4 \times 10^8 \dot{m}^{-1}$ yr. The mean number of active objects at any given time, averaged over the history of the galaxy, required to produce the observed abundance is then

$$\begin{aligned} \langle N \rangle &\approx \frac{10^4 M_\odot}{\epsilon M} \frac{4 \times 10^8 \dot{m}^{-1}}{10^{10}} \\ &\approx 40 \left[\left(\frac{\epsilon}{0.1} \right) \left(\frac{M}{10 M_\odot} \right) \left(\frac{\dot{m}}{10} \right) \right]^{-1} \end{aligned} \quad (9)$$

which may be compared with approximately one hundred appropriate binary X-ray sources occupying the galaxy today. One should also allow for the fact that in typical models of galactic chemical enrichment the current number is likely to be much smaller than the mean, particularly if massive binaries—formed recently from gas which now constitutes only a small fraction of the disk mass—dominate the r-process production. Large apparent fluctuations in r-process production, as indicated for example in fossil Pu in meteorites¹³ and in rare-earth-to-iron abundance ratios in population II stars^{14,15}, argue for a mechanism such as this in which r-process elements are produced by relatively few objects.

We thank M. J. Rees and R. Narayan for valuable conversations. This work was supported by the authors' Alfred P. Sloan Foundation Fellowships and by NASA grant NAGW-763 at the University of Arizona.

Received 30 June; accepted 23 September 1987.

1. Burbidge, E. M., Burbidge, G. R., Fowler, W. A. & Hoyle, F. *Rev. mod. Phys.* **29**, 547-??? (1957).
2. Schramm, D. N. in *Essays in Nuclear Astrophysics* (eds Barnes, C. A., University Press, Clayton, D. D. & Schramm, D. N.) 325-354 (Cambridge University Press, 1982).
3. Mathews, G. J. & Ward, R. A. *Rep. Prog. Phys.* **48**, 1371 (1985).
4. Audouze, J., in *Saas-fee School, The Origin of the Elements* (in the press).
5. Lattimer, J. & Schramm, D. N. *Astrophys. J.* **192**, L145-L147 (1974).
6. Chakrabarti, S. K., Jin, L. & Arnett, W. D. *Astrophys. J.* (submitted).
7. Chakrabarti, S. in *Proc VI Moriond Comp: Accretion in Compact Objects* (in the press).
8. Ramadurai, S. & Rees, M. J. *Mon. Not. R. Astr. Soc.* **215**, 53P-56P (1985).
9. Aharonian, F. A. & Sunyaev, R. A. *Mon. Not. R. Astr. Soc.* **210**, 257-277 (1984).
10. Rees, M. J., Begelman, M. C., Blandford, R. D. & Phinney, E. S. *Nature* **295**, 17-21 (1982).
11. Eardley, D. M., Lightman, A. P., Payne, D. G. & Shapiro, S. L. *Astrophys. J.* **224**, 53-61 (1978).
12. Shapiro, S. L., Lightman, A. P. & Eardley, D. M. *Astrophys. J.* **204**, 187 (1976).
13. Wasserburg, G. & Papanastassiou, D. A. in *Essays in Nuclear Astrophysics* (eds Barnes, C. A., Clayton, D. D., Schramm, D. N.) 77-140 (Cambridge University Press, 1982).
14. Sneden, C. & Parthasarathy, M. *Astrophys. J.* **267**, 757-778 (1983).
15. Sneden, C. & Pilachowski, C. A. *Astrophys. J.* **288**, L55-L58 (1985).
16. Smak, J. *Publ. Astr. Soc. Pacif.* **96**, 5-18 (1984).

Summer dryness in northern mid-latitudes due to increased CO₂

J. F. B. Mitchell & D. A. Warrilow

Meteorological Office, Bracknell, Berkshire RG12 2SZ, UK

Numerical simulations indicate that increases in atmospheric CO₂ (and trace gases) could lead to increased summer dryness of the land surface in northern middle and high latitudes with potentially serious consequences for agriculture (refs 1-5 and M. E. Schlesinger, personal communication). Not all studies support this view, and the validity of the early results has been questioned^{6,7}. We have investigated the dependence of the original finding on the representation of the land surface. We find that the magnitude of the simulated summer drying is dependent on the physical attributes of the soil, and in certain regions can be reduced or even reversed by an alternative, but equally plausible, treatment of run-off over frozen ground.

The experiments were carried out using a version of the Meteorological Office 11-layer atmospheric model coupled to a 50 m 'slab' ocean and a one-dimensional model of sea-ice⁵. The original surface temperature scheme was replaced with a four layer soil temperature model which allows explicitly for the thermal characteristics of different soils. The surface hydrology was amended to include surface run-off and gravitational drainage from the root zone whilst retaining a single soil moisture store (D.A.W., A. B. Sangster and A. Slingo, unpublished Meteorological Office report Met O 20 DCTN 38, May 1986). Run-off increases non-linearly with precipitation (more rapidly so with convective precipitation), and from coarse to fine soils. Drainage increases non-linearly with soil moisture, and decreases from coarse to fine soils. The choice of root depth affects the partitioning between run-off and evaporation. The

value used here (1 m) is more typical of woodland than grassland.

All control simulations were started from a previous control, and anomaly simulations from the corresponding doubled CO₂ simulation, so that the experiments started close to thermal equilibrium. The initial soil moisture in all the integrations, however, was set to that in the original control. Each experiment was run for 6 yr, and the final 5 yr were used in the analysis of the results. (Soil moisture generally requires of the order of one month to reach equilibrium in the tropics, and over 3 months in middle latitudes⁸.) The consistency of the results from year to year indicates that the changes in soil moisture have reached quasi-equilibrium.

The first three experiments used parameters appropriate to clay, medium and sandy soils, respectively. The annual mean and seasonal amplitude of soil moisture in the control simulations decrease from the clay to the sandy soil (Fig. 1a). The annual mean soil moisture is higher for finer soils as they drain more slowly due to their lower hydraulic conductivity. This effect is strongest in winter and spring. In summer, when input is generally high intensity convective precipitation, the finer soils generate greater run-off, leading to reduced moisture absorption, and so appear to dry more rapidly than coarser soils. Values are generally greater than estimates obtained using observed precipitation, an empirical estimate of evaporation and a simple one-layer soil model⁹, largely due to excessive simulated precipitation. In spring, the soil becomes saturated due mainly to snowmelt. The simulated maximum occurs late as the model is too cold in spring.

On increasing CO₂, the soil moisture increases in winter and spring due to enhanced precipitation (Fig. 1b). Snow-melt and hence saturation of the ground and subsequent summer drying occur earlier¹ (Fig. 2), producing a reduction in summer moisture. In spring and summer, convective precipitation increases. This produces an enhancement in surface run-off which is most

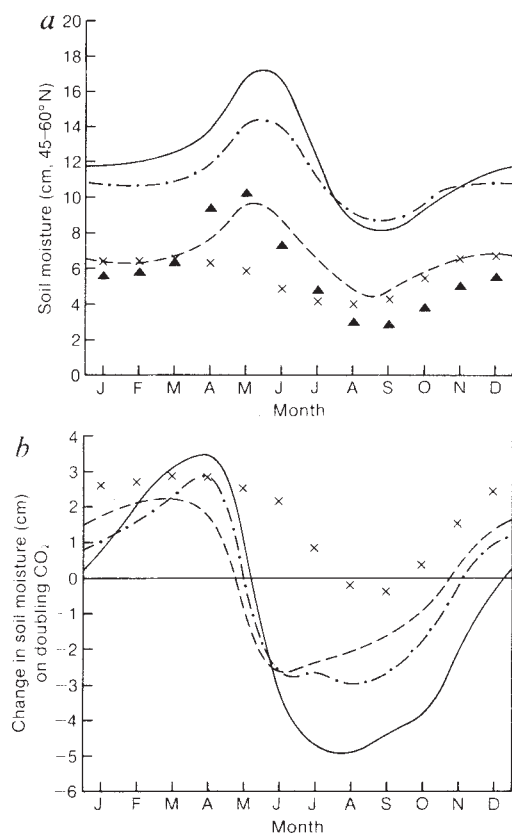


Fig. 1 *a*, Simulated monthly mean soil moisture amounts (cm) averaged over land between 45° and 60° N. Solid line, parameters appropriate to a clay soil; chained dotted line, medium grain soil; dashed line, sandy soil. Crosses, medium soil with 100% run-off over frozen ground (see text). (All results averaged over final 5 years.) The triangles are data derived from observations using empirical estimates of evaporation and a one-layer soil model (ref. 9). *b*, Change in soil moisture (cm) due to doubling atmospheric CO₂ concentrations averaged over land between 45° and 60° N. Symbols as in *a*.

pronounced with the clay soil, so that it dries most on increasing CO₂.

In previous numerical studies of the climatic effects of increased CO₂ it has generally been assumed that all precipitation and snow-melt go to augment soil moisture, unless the soil is saturated^{1,2,4-6}. In practice, infiltration will be limited if the soil is frozen. Furthermore, the release of melt-water to the ground may be abrupt. Successive melting from the snow surface and accumulation or refreezing within the pack increases the density of the snowpack until it becomes 'ripe', after which the structure of the snowpack collapses on further thawing¹⁰⁻¹². The consequent release of melt-water may be sufficiently rapid that little infiltrates the soil, most running off directly. The large but short-lived peak in riverflow in high latitudes following spring snow-melt is consistent with substantial amounts of snow-melt running off directly.

It is not obvious whether or not enhanced run-off over frozen ground would reduce or enhance summer dryness due to increased atmospheric CO₂. Hence we have carried out an additional experiment with a medium soil in which all rain or snow-melt is run off immediately if either of the middle two soil temperature layers (centred at 0.1 and 0.41 m) are sub-freezing. This is an extreme assumption, as some melt-water will penetrate the lower layers if the soil was not saturated when the ground became frozen^{11,12}. This change reduces soil moisture levels in the control simulation. Note that the estimates of actual soil moisture⁹ shown in Fig. 1*a* have been calculated assuming that rain and snow-melt can penetrate frozen ground, and so cannot

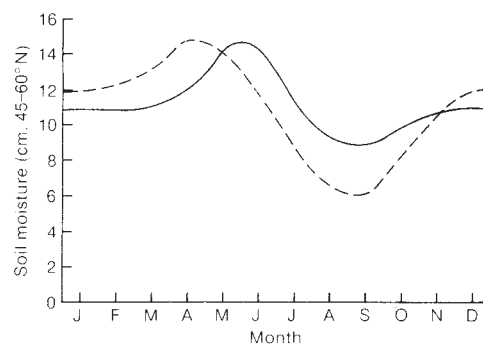


Fig. 2 Simulated monthly mean soil moisture (cm) for a medium soil, averaged over land between 45° and 60° N. Solid line, present CO₂ concentrations; dashed line, doubled CO₂ concentrations.

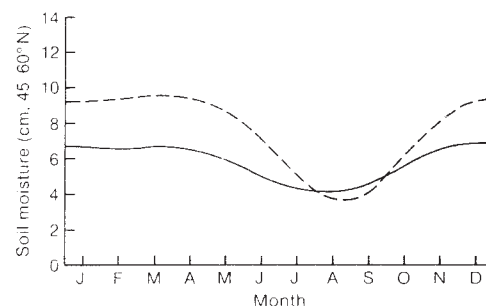


Fig. 3 As Fig. 2, but with 100% run-off when the second and/or third soil layers are frozen.

be used to discriminate between the two alternative formulations for run off. The annual mean run-off increases by less than 10% as increased run-off in spring is compensated by reduced run-off in autumn. This difference is too small to allow us to choose between the methods by comparing annual mean river discharge with simulated river flow calculated by integrating run-off over the catchment area.

With the revised formulation, the soil is no longer saturated in spring (Fig. 1*a*). Thus in the enhanced CO₂ simulation, excess moisture accumulated in winter persists, and although summer drying starts earlier and is faster than in the control, it starts from a higher value (Fig. 3). This contrasts with the original case in which both simulations start from similar saturated values (Fig. 2). Hence, with the revised formulation, the soil does not become drier until late summer (Fig. 1*b*). Note that in the warmer climate, the ground remains unfrozen (and hence able to absorb moisture) over a longer period, contributing to the increase in winter. In another study in which run-off is enhanced over frozen ground^{13,3}, soil moisture does not decrease in mid-latitudes until late summer.

These experiments indicate that the nature of the simulated changes in soil moisture between 45° and 60° N in summer due to increased CO₂ are sensitive to certain aspects of the treatment of the land surface. Other factors including the formulation of evaporation and the effect of vegetation may be important, and are being parameterized in climate models¹⁴⁻¹⁶. Large scale observational studies^{17,18} have been set up to provide data to calibrate such schemes. Increased CO₂ levels may also reduce transpiration and hence alleviate summer dryness¹⁹. These processes will all have to be taken into account before the question as to whether or not increased levels of CO₂ and trace gases will lead to increased summer drying over the northern continents can be answered.

Received 28 July; accepted 2 October 1987.

1. Manabe, S., Wetherald, R. T. & Stouffer, R. J. *Climatic Change* **3**, 347-385 (1981).
2. Mitchell, J. F. B. in *Proc. ISLSCP Conf. Rome, 2-6 Dec 1985*, 171-173 ESA SP-248 (1986).
3. Schlesinger, M. E. & Mitchell, J. F. B. *Rev. Geophys.* **4**, 760-798 (1987).
4. Manabe, S. & Wetherald, R. T. *J. Atmos. Sci.* **44**, 1211-1235 (1987).

5. Wilson, C. A. & Mitchell, J. F. B. *J. geophys. Res.* (in the press).
6. Washington, W. M. & Mechl, G. A. *J. geophys. Res.* **89**, 9475-9503 (1984).
7. MacCracken, M. C., Schlesinger, M. E., Riches, M. R., Manabe, S. & Wetherald, R. T. *Science*, **234**, 659-660 (1986).
8. Yeh, T. C., Wetherald, R. T. & Manabe, S. *Mon. Weath. Rev.* **112**, 474-490 (1984).
9. Willmott, C. J., Rowe, C. M. & Mintz, Y. *J. Clim.* **5**, 589-606 (1985).
10. Dingman, S. L. *et al.* in *Climate, snow cover, micro-climate and hydrology* (eds Brown, J., Miller, P. C., Tiesgen, L. L. & Bunnell, F. L.) Ch. 2 (Dowden, Hutchinson and Ross, Stroudberg, Pennsylvania, 1980).
11. Ward, R. in *Floods. A geographical perspective*, Chs 2, 3 (Macmillan, London, 1978).
12. Raudkivi, A. J. in *Hydrology. An Advanced Introduction to Hydrological Processes and Modelling*, Ch. 6 (Pergamon, Oxford, 1979).
13. Hansen, J. *et al.* *Geophys. Monogr.* **29**, 130-163 (1984).
14. Sellers, P. J., Mintz, Y., Sud, Y. C. & Dalcher, A. *J. atmos. Sci.* **43**, 505-531 (1986).
15. Dickinson, R. E. in *Climate Sensitivity* (eds Hansen, J. & Takahashi, T.) Geophysical Monograph 29, 130-163 (American Geophysical Union, Washington, DC, 1984).
16. Warrilow, D. A. in *Proc. ISLSCP Conf. Rome, 2-6 Dec. 1985*, 159-166, ESA SP-248 (1986).
17. *Development of the implementation plan for the International Satellite land-surface climatology project (ISLSCP) Phase I* (eds Bolle, H. J. & Rasool, S. I.) WCP-94 (World Meteorological Organization, Geneva, 1985).
18. *Report of the second meeting of the scientific steering group on land surface processes and climate* (Toulouse, France, 21-26 May 1986). WCP-126 (World Meteorological Organization, Geneva, 1987).
19. Cure, J. D. & Acock, B. *Agriculture and Forest Meteorology* **38**, 127-145 (1986).

Measurements of methanesulphonic acid in Antarctic ice

C. Saigne & M. Legrand

Laboratoire de Glaciologie et Géophysique de l'Environnement,
BP 96 38402 St Martin d'Hères Cedex, France

Dimethylsulphide (DMS), mainly produced by marine biogenic activity, plays an important role in the atmospheric sulphur budget¹. Methanesulphonic acid (MSA) and sulphur dioxide (SO₂, hereafter converted into non-sea-salt (n.s.s.) sulphate) are the main oxidation products of DMS². As opposed to n.s.s. sulphate, which has other sources (for example, volcanoes, terrestrial sulphate), MSA represents an unequivocal indicator of marine biogenic activity. MSA has previously been investigated in marine aerosols at low and mid-latitudes³⁻⁵ as well as in precipitations and polar ice^{5,6}; here we present 54 MSA measurements made in Antarctic ice. Coastal area precipitations exhibit unexpectedly high (up to 100%) MSA/n.s.s. SO₄²⁻ weight ratios (*r*) compared with values commonly observed in mid-latitude marine atmospheres³. At higher elevations (2,000 m) the *r* values suggest a marine biogenic input of more global significance. The MSA concentrations (several p.p.b.) confirm that the n.s.s. sulphate in Antarctic ice is mainly derived from marine biogenic activity. During the last ice age, MSA contents were 2-5 times higher than today. This study of high-latitude precipitations demonstrates the feasibility of reconstructing past marine biogenic activity of global significance.

Nineteen sections were selected from the length of the 900-m-deep Dome C and thirty-five sections from the 300-m-deep D10 ice cores respectively. Each 11-cm-long section (covering ~5 yr of precipitation at Dome C and ~0.5 yr at D10) was first washed with ultrapure water, following the decontamination procedure designed for inorganic trace species measurements in ice⁷. But, because other organic compounds (formic and acetic acids, not discussed here⁸), were also measured in these samples, the subsequent procedure was followed. The pre-washed samples were placed in an airtight glass device and washed with ultrapure water under pure nitrogen circulation. Samples were then analysed by ion chromatography. An MSA measurement requires 7 min and a 5-ml sample is sufficient for a detection limit of 0.16 ng g⁻¹ (ref. 9). The analytical precision is typically ±10% (at a 95% confidence level). A more detailed description of the method and analytical tests are reported elsewhere⁹. For each sample, we also measured Cl⁻ and SO₄²⁻.

Both Dome C (east Antarctica, 3,250-m elevation) and D10 (Terre Adélie, 270-m elevation) deep ice cores cover the last several thousand years including the last great climatic change

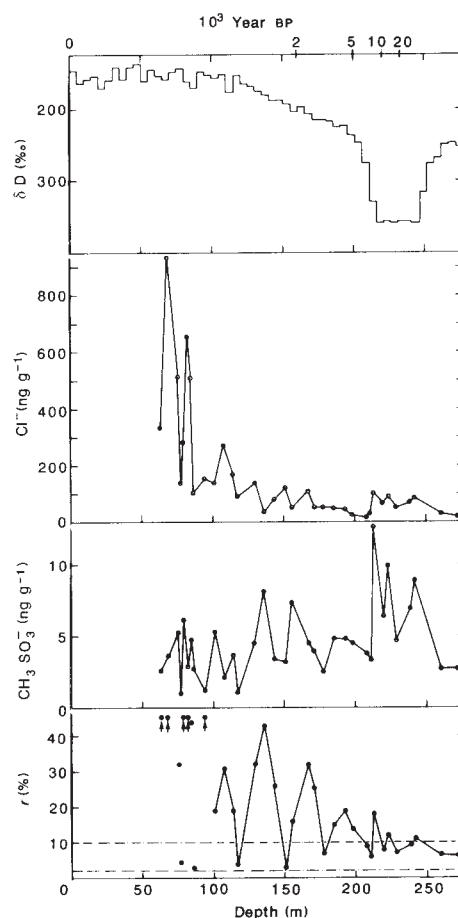
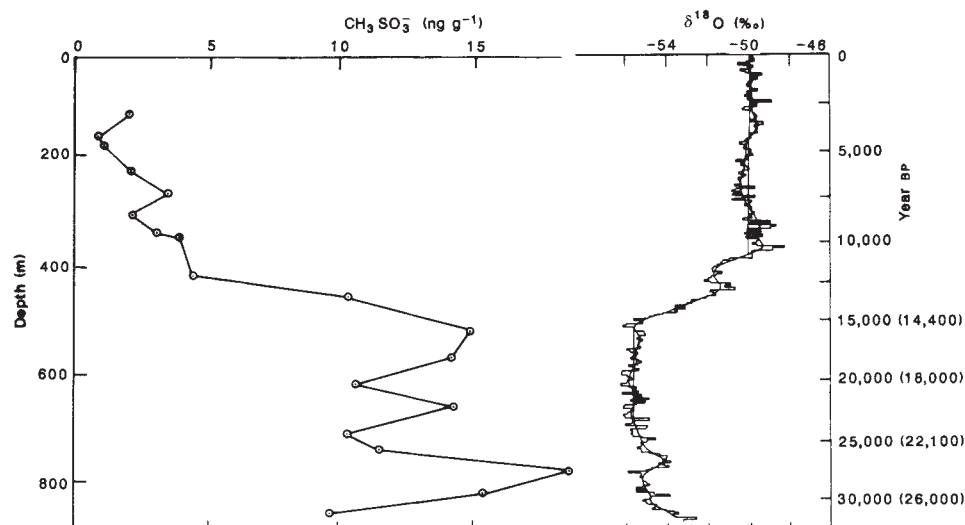


Fig. 1 Cl⁻, MSA (CH₃SO₃⁻) and MSA/n.s.s. SO₄²⁻ (*r*) down the D10 ice core; horizontal dashed lines refer to marine values observed by Saltzman *et al.*³. Interpretation of the isotope profile (δD, ‰) is complicated by an important ice-flow phenomenon. Indeed, δD is a function of temperature modulated by both past climatic changes and the origin of the ice (elevation). As suggested by Raynaud *et al.*¹⁷, the ice is of local origin down to a depth of 100 m. The subsequent δD decrease reflects a higher ice origin site (~2,000 m). Down to 210-m depth, the ice covers the last 10,000 yr (the Holocene). The abrupt δD decrease at ~210 m is attributed to the climatic change occurring ~15,000 yr BP. The dating and origin of the ice below 250 m remains unclear and is not discussed here.

(~15,000 yr BP). Moreover, at D10 (drilled at 5 km from the coast) isotope variations (δD, Fig. 1) with depth are related to both past climatic change and the geographical origin of the ice (due to an important ice-flow phenomenon). The study of this core therefore allows us to examine spatial ice chemistry variations and variations with climate.

Because the Cl/Na weight ratio of ice is very close to the reference bulk seawater ratio (1.8) throughout the length of the D10 core¹⁰ both chloride and sodium can be considered as conservative tracers of sea-salt content. We chose to use Cl⁻ for practical reasons (measured with sulphate). Down to a depth of 100 m high chloride values are observed (406 ± 261 ng g⁻¹, Table 1). Then the chloride content decreases rapidly, levelling off between 175 and 210 m (35 ± 13 ng g⁻¹). Due to the ice flow phenomenon this chloride profile reflects the deposition of the coarsest sea-salt particles very near the coast. For MSA, the mean content is not significantly different for depth intervals 60-100 and 100-210 m (3.4 ± 1.7 and 4.2 ± 1.7 ng g⁻¹ respectively). These different patterns of Cl⁻ and MSA profiles suggest that MSA and sea-salt are not deposited by the same process. Therefore, our data confirms observations made by Saltzman *et*

Fig. 2 MSA (CH_3SO_3^-) down the Dome C core; the isotope profile ($\delta^{18}\text{O}$) indicate the Holocene and the end of the last ice age¹⁸. The dating of this core is plotted at the top¹⁸.



al^3 from a study of particle-size distribution in marine aerosol. They found that MSA was concentrated primarily in the smaller particles, and not in the coarser sea-salt particles. It seems that although MSA is formed by a homogeneous gas-phase reaction, it is unlikely to remain in the gas phase due to its extreme solubility¹¹. It is probably extensively scavenged by water so that it does not react significantly with the coarser alkaline sea-salt particles.

DMS can react with various oxidizing species including hydroxyl (OH) and nitrate (NO_3) radicals¹² to produce mainly SO_2 and MSA. But, reaction sequences and relative amounts of products are still quite uncertain. Indeed from experimental studies, Hatakeyama *et al.*¹³ predicted that the DMS oxidation reaction provides 50 and 20% of MSA and SO_2 respectively whereas Grosjean and Lewis² suggested that SO_2 is the main product. In samples from remote marine areas, the calculated MSA/n.s.s. SO_4^{2-} weight ratio values (r) vary between 2 and 10%³. To explain these low r values in the real atmosphere, Saltzman *et al.* proposed that MSA is unstable in aerosols and cloud droplets because of a conversion with OH radicals into n.s.s. sulphate³. Down to 100 m, our samples show very scattered values with several points $>100\%$ (Fig. 1). Below 100 m depth, the r values decrease gradually, reaching $\sim 10\%$ at 175-m depth and then becoming consistent with Saltzman's values along the remaining part of the core while no corresponding MSA content trend is observed. Down to 100 m, as suggested by isotope and chloride profiles (Fig. 1), the high r values correspond to local precipitations which reflect the marine atmosphere of these high

latitudes. When leaving the coastal area for the high plateau, the presence of very high pressure causes the influence of oceanic air masses to become weaker. In these areas, the atmosphere is considered as a 'background atmosphere' mainly supplied by long-range transport of impurities produced far from Antarctic locations¹⁴. Therefore, as an explanation for the observed spatial variations, we would like to suggest that the high r values in coastal ice reflect either a relatively rich MSA production in the sub-Antarctic ocean or a lower OH content in cloud droplets at these high latitudes. Therefore when moving inland, the r values are more similar to Saltzman's values in relation to marine input from remote marine areas.

At D10 in the ice deposited at the end of the last ice age (between 210 and 250-m depth, Fig. 1) we observe chloride contents double that found during the Holocene. This result confirms previous studies of several other Antarctic ice cores which have indicated a large increase of marine input during the ice age mainly due to strengthened source-area production, enhanced poleward transport and to a lesser extent, a reduced accumulation rate of snow^{15,16}. We observe a twofold increase in MSA and n.s.s. SO_4^{2-} during the ice age. This increase can reflect changes of global (oceanic productivity, transport efficiency) and local (accumulation rate of snow) significance. From this rather weak change it is difficult to evaluate the relative contribution of these parameters due to a lack of knowledge of the past accumulation rate of snow at this location¹⁷. At Dome C (Fig. 2) we observe a fivefold increase in MSA content during the ice age while the accumulation rate of snow was a factor of 1.7 lower¹⁸. Hence, even if a part of the increase reflects the reduced accumulation rate of snow, the results suggest an increase of MSA input reflecting a higher marine biogenic activity of global significance and enhanced transport between source areas and Antarctica. Although it is still impossible to evaluate the relative contribution of these two parameters, the data agrees with a previous study of marine palaeoproductivity indicating a threefold increase in biogenic productivity during the ice age in upwelling areas¹⁹.

Financial support was provided by the French Centre National de la Recherche Scientifique (CNRS). Field work was supported by Terres Australes et Antarctiques Françaises, Expéditions Polaires Françaises and the NSF.

Table 1 D10 ice core: mean concentration and standard deviation (1σ) of all measured parameters for four isotopic stages

Depth (m)	Cl^- (ng g^{-1})	MSA (ng g^{-1})	n.s.s. SO_4^{2-} (ng g^{-1})	r (%)
60–100	402 ± 261	3.4 ± 1.7	$-13.5^* - 89.0$	3–100
100–210	86 ± 64	4.2 ± 1.7	31.0 ± 23.0	18.9 ± 11.0
(175–210)	(35 ± 13)	(3.9 ± 0.8)	(37 ± 9)	(12.0 ± 4.7)
210–250	77 ± 17	8.2 ± 2.6	75 ± 7	11.0 ± 3.5
Ice age/Holocene	2.2	2.1	2.03	1.1

The 60–100 m interval refers to local precipitation, 100–210 to the ice flow phenomenon and 210–250 to the ice age (see Fig. 1). To minimize the ice flow problem the ratio between ice age and Holocene values was calculated using the 175–210 m interval for the Holocene. This ratio therefore reflects a change at $\sim 15,000$ yr BP at $\sim 2,000$ -m elevation.

* Negative data obtained for n.s.s. sulphate may be explained by the high uncertainty of measurements of this compound: $\pm 8 \text{ ng g}^{-1}$.

Received 29 April; accepted 29 August 1987.

1. Andreae, M. O. & Raemdonck, H. *Science* **221**, 744–747 (1983).
2. Grosjean, D. & Lewis, R. *Geophys. Res. Lett.* **9**, 1203–1206 (1982).
3. Saltzman, E. S., Savoie, D. L., Zika, R. G. & Prospero, J. M. *J. geophys. Res.* **88**, 10897–10902 (1983).
4. Saltzman, E. S., Savoie, D. L., Prospero, J. M. & Zika, R. G. *Geophys. Res. Lett.* **12**, 437–440 (1985).

5. Ayers, G. P., Ivey, J. P. & Goodman, H. S. *J. Atmos. Chem.* **4**, 173-185 (1986).
6. Ivey, J. P., Davies, D. M., Morgan, V. & Ayers, G. P. *Tellus* **38B**, 375-379 (1986).
7. Legrand, M., De Angelis, M. & Delmas, R. *J. Analyt. chim. Acta* **156**, 181-182 (1984).
8. Legrand, M. & Saigne, C. *Atmos. Envir.* (submitted).
9. Saigne, C., Kirchner, S. & Legrand, M. *Analyt. chim. Acta* (in the press).
10. Legrand, M. thèse d'état, Univ. Scientifique et Médicale de Grenoble (France) *Laboratoire de Glaciologie du CNRS* (1985).
11. Clegg, S. L. & Brimblecombe, P. *Envir. Tech. Lett.* **6**, 269-278 (1985).
12. Atkinson, R., Pitts, J. N. Jr. & Aschmann, S. M. *J. phys. Chem.* **88**, 1584-1587 (1984).
13. Hatakeyama, S., Izumi, K. & Akimoto, H. *Atmos. Envir.* **19**, 135-141 (1985).
14. Cunningham, W. C. & Zoller, W. H. *J. Aerosol Sci.* **12**, 367-384 (1981).
15. Petit, J. R., Briat, M. & Royer, A. *Nature* **293**, 391-394 (1981).
16. Palais, J. & Legrand, M. *J. Geophys. Res.* **90**, 1143-1154 (1985).
17. Raynaud, D., Lorius, C., Budd, W. F. & Young, N. W. *J. Glaciol.* **24**, 103-114 (1979).
18. Lorius, C., Merlivat, L., Jouzel, J. & Pourchet, M. *Nature* **280**, 644-648 (1979).
19. Sarthein, M., Winn, K. & Zahn, R. *Abrupt Climatic Changes: Evidence and Implications* (eds Berger, W. H. & Labeyrie) (Reidel, Dordrecht, in the press).

Sulphur isotope heterogeneity in the mantle from ion microprobe measurements of sulphide inclusions in diamonds

M. Chaussidon, F. Albarède & S. M. F. Sheppard

Centre de Recherches Pétrographiques et Géochimiques et
ENS Géologie BP 20, 54501 Vandœuvre Cedex, France

The sulphur isotope composition of the mantle has been generally considered to be $0 \pm 2\%$ ¹. There is, however, considerable uncertainty in the mean $\delta^{34}\text{S}$ value of the mantle and in its range because there have been very few analyses of unaltered sulphide phases of direct mantle origin¹⁻⁴ and $\delta^{34}\text{S}$ values of sulphides from large mafic intrusions range from -9 to $+17\%$ ⁵. We report here ion microprobe $\delta^{34}\text{S}$ determinations from two mantle derived sulphide inclusions ($\approx 100\ \mu\text{m}$) preserved in a diamond ($\delta^{13}\text{C} = -4.25\%$) of peridotitic(?) paragenesis from Premier (South Africa) and in four diamonds of eclogitic paragenesis from Orapa (Botswana). Three octahedral sulphides have lower $\delta^{34}\text{S}$ values ($+2.3 \pm 1.4\%$) than three platelet-like sulphides ($+8.2 \pm 0.9\%$). This evidence for mantle heterogeneity for sulphur is interpreted in terms of recycling of crustal material into the eclogitic region of diamond growth in the mantle.

Two octahedral sulphide inclusions from the same diamond from Premier and one octahedral plus three platelet-shaped inclusions from four different diamonds from Orapa were analysed with both electron and ion microprobes (Table 1). The inclusions are about $100\ \mu\text{m}$ in size and are considered primary for the following reasons. No crack was observed in the

Table 2 Instrumental mass fractionation factors of sulphides standards for $\delta^{34}\text{S}$ measurements using an IMS3F ion probe

Mineral	Instrumental mass fractionation (‰)	$\Delta^*_{\text{pyrrh-min}}$
Pyrrhotite	-64.4	0
Pentlandite	-48.3	-16.1
Pyrite	-69.0	+4.6
Chalcocopyrite	-69.1	+4.7
Arsenopyrite	-66.5	+2.1
Galena	-55.3	-9.1
Cinnabar	-66.5	+2.1

* The parameter $\Delta^*_{\text{pyrrh-min}}$ (difference of instrumental mass fractionation between pyrrhotite and mineral) is included as it is supposed to be less dependent on the analytical conditions than instrumental mass fractionation.

diamonds around the inclusions. The form of the inclusion was imposed by the diamond during its growth, perhaps at different stages for the octahedral inclusions and for the platelet inclusions which are located along crystallographic planes of the diamond (J. W. Harris, personal communication). The sulphides (Table 1) are a pentlandite and five pyrrhotites with very thin exsolved pentlandite ($\approx 1\ \mu\text{m}$), sometimes with exsolved chalcocopyrite (several μm) in the periphery as previously described for sulphide paragenesis of magmatic origin^{4,6-10}. The compositions of the pyrrhotites plot in the Fe, Ni, Cu, S monosulphide-solid-solution series (Mss)¹¹ and are consistent with being high temperature ($\approx 1,000\ ^\circ\text{C}$) phases which underwent closed system exsolution of chalcocopyrite at around $850\ ^\circ\text{C}$ and pentlandite at about $610\ ^\circ\text{C}$ ^{4,12}. The four Orapa diamonds contained silicate inclusions of an eclogitic suite (J. W. Harris, personal communication); the Fe/Fe + Ni ratios of the sulphide inclusions are also characteristic of eclogitic types¹³. The diamond from Premier did not contain silicate inclusion and its $\delta^{13}\text{C}$ value of -4.25% (Table 1) cannot be used to classify the inclusion suite^{14,15}. The Fe/Fe + Ni ratio of the sulphide inclusion as a whole cannot be determined because the inclusion broke into two pieces, one of pyrrhotite and the other of pentlandite. Nevertheless the initial inclusion must have had a relatively high Ni/Fe ratio because the two fragments are approximately of the same size; this inclusion may therefore be of peridotitic type¹³.

The sulphur isotope composition of the inclusions was determined using a modified CAMECA IMS 3F ion microprobe¹⁶. The inclusions were mounted in epoxy, polished and coated

Table 1 Isotopic analyses of sulphide inclusions within diamonds from Premier and Orapa

Inclusion	Shape and size	Ni/Fe	Cu/Fe	Electron probe		Fe/Fe + S	$\delta^{34}\text{S}^\dagger$ ‰	Ion probe Pentlandite in the analysed area (%)	
				As/Fe	Fe/Fe + Ni				
PREMIER ($\delta^{13}\text{C} = -4.25 \pm 0.07\%$ (2 analyses), peridotitic (?) paragenesis):									
7a: pyrrhotite	Octahedral	0.076	0.007	0.001	0.929	0.584	+1.7	8.7	
7b: pentlandite	Octahedral	ND	ND	ND	ND	ND	+1.1	≈ 100	
ORAPA (eclogitic paragenesis):									
OR 401: pyrrhotite	Octahedral	0.004	0	0.001	0.996	0.599	+2.1	0	
	$300 \times 70 \times 70\ \mu\text{m}$						+4.3*	0	
OR 403: pyrrhotite	Platelet	0.021	0	0.001	0.979	0.589	+8.3	4.5	
	$100 \times 70 \times 30\ \mu\text{m}$						+7.1*	3.8	
OR 404: pyrrhotite	Platelet	0.017	0.015	0.001	0.984	0.597	+8.5	2	
	$100 \times 100 \times 30\ \mu\text{m}$						+9.5*	0	
OR 418: pyrrhotite	Platelet	0.023	0.001	0.001	0.978	0.592	+7.8	1.8	
	$180 \times 180 \times 20\ \mu\text{m}$								

ND, not determined.

* New analysis after repolishing of the sample.

$\dagger \delta^{34}\text{S} = 1,000 \times ((^{34}\text{S}/^{32}\text{S})_{\text{sample}} / (^{34}\text{S}/^{32}\text{S})_{\text{CDT}} - 1)$. CDT, Canyon Diablo Troilite standard.

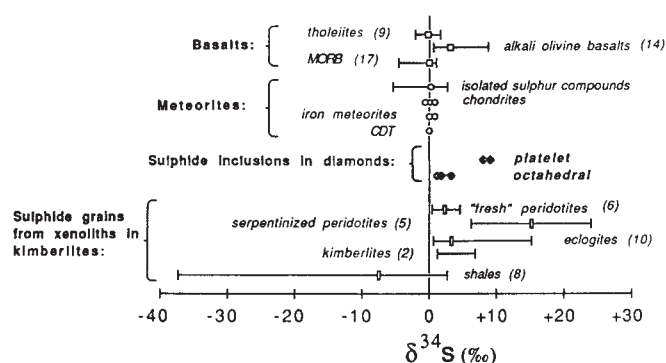


Fig. 1 Variations of the sulphur isotope composition of sulphide inclusions in diamonds and other mantle derived rocks. Numbers of samples used for calculation of the mean shown in parentheses. Data on basalts from refs 2, 22; on meteorites from refs 19, 20; diamond inclusion data from this study; and xenolith data from refs 3, 4.

with gold. A negative primary beam of oxygen rastered over 100 μm with an analysed area of 60 μm in diameter and a secondary positive beam were used without energy filtering. The mass resolution was 3,000, or 4,500 when interferences with hydrides were observed. Data were acquired using a Solartron 7060 digital voltmeter. The instrument was calibrated for mass fractionation (Table 2) and linearity verified on pyrite for $\delta^{34}\text{S}$ between -25% to $+20\%$ using a set of mineral standards analysed both chemically and isotopically. Because the instrumental mass fractionation is sensitive to the nature of the sulphide phase, the Fe, Cu and Ni abundances in the area analysed for sulphur isotopes were routinely monitored. The reproducibility was generally better than $\pm 1\%$ while the within-run statistics is about $\pm 0.7\%$.

The $\delta^{34}\text{S}$ (CDT) of the sulphides (Table 1) can be divided into two groups: the three octahedral inclusions with a mean value of $+2.3 \pm 1.4\%$ and the three platelet inclusions with a mean value of $+8.2 \pm 0.9\%$. There is no apparent correlation between $\delta^{34}\text{S}$ and either the Ni or Cu contents of the inclusions. These $\delta^{34}\text{S}$ values are considered to record the sulphur isotope compositions of the immiscible melts that were trapped during the growth of the diamonds in the mantle. They were therefore armoured from possible subsequent alteration processes.

The $\delta^{34}\text{S}$ of the sulphide inclusions not only demonstrate that the mantle is not homogeneous for sulphur but that variations of $\sim 5\%$ occur in diamonds from a single kimberlite. Because of the generally accepted long and complex growth history of diamonds^{15,17,18}, the octahedral and platelet inclusions probably represent samples of different parts of the mantle and at different times.

The commonly proposed $\delta^{34}\text{S}$ value of $0 \pm 2\%$ for the mantle is based on (1) the extremely restricted range of values for most sulphides from iron and chondritic meteorites (Fig. 1)^{19,20}; (2) early analyses of sulphides from major mafic intrusions formed by partial melting of the mantle²¹; (3) data on total sulphur in continental basalts and MORB which have not undergone sub-aerial degassing^{2,22}; and (4) some of the analyses of sulphides, which could be primary, from mantle derived xenoliths^{3,4}. These results as a whole are in broad agreement with our $\delta^{34}\text{S}$ values of octahedral sulphide inclusions. Although the platelet sulphide inclusions are markedly ^{34}S enriched relative to values for MORB or meteorites, several of the pyrrhotites in eclogite xenoliths from kimberlites have similar or even higher $\delta^{34}\text{S}$ values (Fig. 1).

Assuming that the $\delta^{34}\text{S}$ value of the Earth is zero, like that of the meteorites, certain regions of the upper mantle have become ^{34}S enriched. Such regions could have preferentially either lost ^{32}S or gained ^{34}S . There are three possibilities. (1)

^{32}S has been lost to a deeper hidden reservoir whose $\delta^{34}\text{S}$ value has become less than zero; the $\delta^{34}\text{S}$ already reported for a sulphide globule within a high temperature (1,500 $^{\circ}\text{C}$) clinopyroxene megacryst from the Uintjesberg kimberlite (South Africa) of -2.2% ²³ could be a possible signature of such a reservoir. (2) A ^{32}S -enriched volatile phase like H_2S for an oxygen pressure below 10^{-8} atm (ref. 17) produced by a reaction such as $2\text{FeS} + \text{CH}_4 \rightarrow 2\text{H}_2\text{S} + 2\text{Fe} + \text{C}$, could be lost from the sulphur saturated melts in which the diamonds grew: however according to a Rayleigh distillation model, degassing of 99% of the initial sulphur with an extreme gas-melt fractionation at $\sim 1,000$ $^{\circ}\text{C}$ of -0.5% for H_2S ^{24,25} can only account for an increase of 2.3% in the isotopic composition of the residual sulphide. (3) A ^{34}S -enriched component has been added; subduction of seawater-hydrothermally altered oceanic crust which has thus become enriched in ^{34}S by up to $+10\%$ ²⁶, is a suitable source. Of these possibilities, contamination of the upper mantle by subducted oceanic crust appears to be the only realistic process based on the available isotopic results and current knowledge of the fractionation factors. A similar process has been suggested to account for the $\delta^{15}\text{N}$ values of diamonds from Zaire²⁷ and also to modify the $\delta^{34}\text{S}$ of rocks from the Marianas island arc²⁸.

This mantle contamination explanation imposes certain limits on the age of the inclusions with $\delta^{34}\text{S} \approx +8\%$. Although the sulphur isotope compositions of the principal terrestrial reservoirs and their evolution are not well constrained¹, crustal sulphides and sulphates probably did not deviate by more than $\pm 5\%$ from 0 before about 2,700 Myr ago²⁹. Only a short time is needed for the incorporation at depth of oceanic crust: for example, 215 Myr for 150 km depth and 1 mm per year horizontal speed of the plate. This implies that these inclusions are younger than about 2,500 Myr.

Although the data base is still limited, the age determinations of silicate inclusions in diamonds, 3,300 Myr for peridotitic inclusions from Kimberley and Finsch³⁰ and 1,150–1,580 Myr for eclogitic inclusions from Premier and the Argyle lamproite³¹ suggest that eclogitic diamonds could be younger than peridotitic ones. However, neither age nor $\delta^{34}\text{S}$ data on eclogitic and peridotitic inclusions from the same kimberlite are yet available. Although the ages of inclusions in diamond from the Cretaceous Orapa kimberlite³² (93.1 Myr) are not yet known, the suggested younger age of the eclogitic inclusions are not in contradiction with the interpretation of $\delta^{34}\text{S} > +5\%$ in terms of mantle contamination.

The significance of the most ^{34}S -depleted ($+3.2\%$) inclusion (OR 401), which is octahedral, from Orapa is not fully understood. The fact that $\delta^{13}\text{C}$ data have shown very complex growth histories for diamonds^{18,33} make the growth of diamond in different environments possible. Thus this diamond could have trapped this inclusion in a different eclogitic environment from that of the platelet inclusions. Alternatively, even though eclogitic and peridotitic suite silicate inclusions have not yet been reported from the same diamond, at Orapa, no sharp compositional gap has been observed between the peridotitic and eclogite suite of silicates³⁴. The diamond (OR 401) could then have started to grow in a peridotitic environment around a sulphide globule, which crystallized as an octahedral inclusion, and completed its growth in an eclogitic environment.

These first results on sulphide inclusions in diamonds demonstrate the potential of the ion probe technique in tackling otherwise inaccessible problems and suggest that future measurements on previously well-documented diamonds will provide useful information. Studies are continuing to relate $\delta^{34}\text{S}$ values with $\delta^{13}\text{C}$ data, nitrogen values and silicate inclusion type.

We thank J. W. Harris and C. B. Smith for providing the samples and beneficial discussions, E. Deloule for advice on ion probe analyses, J. C. Demange and B. Jacquier for technical assistance and S. E. Haggerty for comments. This study was supported by ATP "Distribution des éléments...", CNRS.

Received 4 August; accepted 5 October 1987.

1. Nielsen, J. in *Lectures in Isotopes Geology*, (eds Jager, E. & Hunziker, J. C.) 283-312 (Springer, Berlin, 1979).
2. Schneider, A. *Contr. Mineral. Petrol.* **25**, 95-124 (1970).
3. Grinenko, V. A., Dmitriev, L. V., Migdisov, A. A. & Sharas'kin, A. Ya. *Geochem. Int.* **12**, 132-137 (1975).
4. Tsai, H. M., Shieh, Y. & Meyer, H. O. A. in *The Mantle Samples: Inclusions in Kimberlites and Other Volcanics* Vol. 2 (eds Boyd, F. R. & Meyer, H. O. A.) 87-103 (American Geophysical Union, Washington DC, 1979).
5. Ohmoto, H. in *Stable Isotopes in High Temperature Geological Processes. Reviews in Mineralogy* Vol. 16 (ed. Ribbe, P. H.) 491-556 (Mineralogical Society of America, Washington DC, 1987).
6. Skinner, B. J. & Peck, D. L. *Econ. Geol. Monogr.* **4**, 310-322 (1969).
7. Harris, J. W. *Contr. Mineral. Petrol.* **35**, 22-33 (1972).
8. Kanehira, K., Yui, S., Sakai, H. & Sasaki, A. *Geochem. J.* **7**, 89-96 (1973).
9. Mathez, E. A. *J. geophys. Res.* **81**, 4269-4276 (1976).
10. Lorand, J.-P. & Conqu  r  , F. *Bull. Mineral.* **106**, 585-605 (1983).
11. Craig, J. R. & Kuller  d, G. *Econ. Geol. Monogr.* **4**, 344-358 (1969).
12. Lorand, J.-P. *Bull. Soc. g  ol. France* **8**, III, 643-656 (1987).
13. Efimova, E. S., Sobolev, N. V. & Pospelova, S. N. *Lett. All Union Mineral. Soc.* **3**, 300-310 (1983).
14. Deines, P., Gurney, J. J. & Harris, J. W. *Geochim. cosmochim. Acta* **48**, 325-342 (1984).

15. Meyer, H. O. A. *Am. Miner.* **70**, 344-355 (1985).
16. Deloule, E., Allegre, C. & Doe, B. *Econ. Geol.* **81**, 1307-1321 (1986).
17. Haggerty, S. E. *Nature* **320**, 34-38 (1986).
18. Swart, P. K., Pillinger, C. T., Milledge, H. J. & Seal, M. *Nature* **303**, 793-795 (1983).
19. Thode, H. G., Monster, J. & Dunford, H. B. *Geochim. cosmochim. Acta* **25**, 159-174 (1961).
20. Kaplan, I. R. & Hulston, J. R. *Geochim. cosmochim. Acta* **30**, 479-496 (1966).
21. Shima, M., Gross, W. H. & Thode, H. G. *J. geophys. Res.* **68**, 2835-2847 (1963).
22. Sakai, H., Des Marais, D. J., Ueda, A. & Moore, J. G. *Geochim. cosmochim. Acta* **48**, 2433-2441 (1984).
23. Chaussidon, M., Albarede, F. & Sheppard, S. M. F. *Eos* **68**, 440 (1987).
24. Sakai, H. *Geochim. cosmochim. Acta* **12**, 150-169 (1957).
25. Ohmoto, H. and Rye, R. O. in *Geochemistry of Hydrothermal Ore Deposits* 2nd ed (ed. Barnes, H. L.) 509-567 (Wiley, New York, 1979).
26. Albarede, F. & Michard, A. *Chem. Geol.* **57**, 1-15 (1986).
27. Javoy, M., Pineau, F. & Demaiffe, D. *Earth planet. Sci. Lett.* **68**, 399-412 (1984).
28. Woodhead, J. D., Harmon, R. S. & Fraser, D. G. *Earth planet. Sci. Lett.* **83**, 39-52 (1987).
29. Schidlowski, M., Hayes, J. M. & Kaplan, I. R. in *The Earth's Earliest Biosphere. Its Origin and Evolution* (ed. Schopf, J. W.) 149-186 (1983).
30. Richardson, S. H., Gurney, J. J., Erlank, A. J. & Harris, J. W. *Nature* **310**, 198-202 (1984).
31. Richardson, S. H. *Nature* **322**, 623-626 (1986).
32. Davis, G. L. *2nd Int. Kimb. Conf. Santa Fe, Abstr.* (1977).
33. Deines, P., Harris, J. W. & Gurney, J. J. *Geochim. cosmochim. Acta* **51**, 1227-1243 (1987).
34. Gurney, J. J., Harris, J. W. & Rickard, R. S. in *Kimberlites II: The Mantle and Crust-Mantle Relationships* (ed. Kornprobst, J.) 3-9 (Elsevier, Amsterdam, 1984).

Prediction of coastal sediment stability from photopigment content of mats of purple sulphur bacteria

Jonathan Grant* and Giselher Gust†

* Department of Oceanography, Dalhousie University, Halifax, Nova Scotia, B3H 4J1 Canada

† Department of Marine Science, University of South Florida, St Petersburg, Florida 33701, USA

The threshold of sediment erosion by currents and waves (critical friction velocity, u_{crit}) is an important predictive tool in sedimentology and coastal engineering^{1,2} owing to its use in entrainment and transport rate equations. The measurement of sediment transport is relevant to many concerns including dispersal of pollutants and dredge spoil, harbour and beach maintenance, geochemical fluxes and animal-sediment relations³. In marine sediments, biotic processes alter erosion criteria⁴, for instance through binding and stabilization of the sediment surface by the extracellular products of diatoms, bacteria and other benthic organisms^{5,6}. Despite the widespread occurrence of benthic microbial mats and attendant sediment binding, prediction of erosion thresholds and transport rates in natural marine sediments has incorporated biotic variables only qualitatively^{5,7}. We used a laboratory flume to erode intact sediment cores containing undisturbed mats of phototrophic purple sulphur bacteria. The pigment content of bacterial films (bacteriochlorophyll *a*) was positively correlated with u_{crit} , increasing the erosion threshold as much as fivefold compared with sterile control cores. We demonstrate the profound influence of benthic microbiota on prediction of erosion in marine sediments.

Mats of purple sulphur bacteria (Chromatiaceae) are widespread in sulphide-rich, well-lit habitats⁸. In south Florida these mats are visible on many beaches. At our study site (Pinellas Point, Tampa Bay, Florida) a zone of purple mats 1-2 m wide occurred on the top of an occasionally exposed sand bar (mean grain diameter 189 μm , 20-25 $^{\circ}\text{C}$, 28-29%). Within the areas of purple bacteria, cm-scale variation gave the sediment a mottled appearance (Fig. 1a). The mats were interspersed with patches of white ('clear') surface sediment containing fewer purple sulphur bacteria and no visible mats.

Vertically, laminae of purple and clear sediment layers (2-mm thickness) were underlain by a greenish layer of filamentous cyanobacteria (2-mm thickness) and then by gray-black reducing sediments and/or more purple bacterial bands. Purple sand grains from the surface showed attached purple cell clumps characteristic of many Chromatiaceae⁹ (Fig. 1b). Extracellular slime associated with these cell clumps¹⁰ linked sand grains

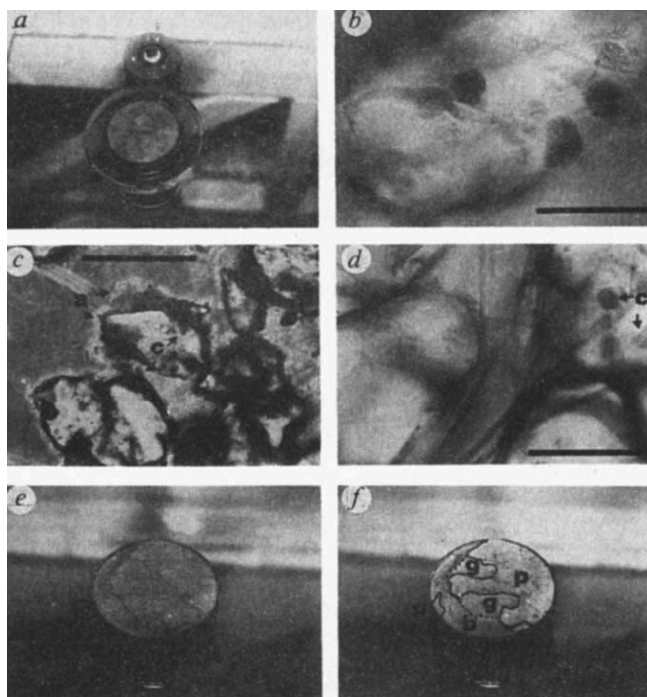


Fig. 1 Laminated microbial mats from Pinellas Point, Tampa Bay, Florida. *a*, Intact sediment core (7.5-cm diameter) emplaced in the flume channel, the sediment surface flush with the flume bed. Note the mottled appearance due to the presence of purple sulphur bacteria (dark regions). The circular object (1.2-cm diameter) adjacent to the core is the flush-mounted, hot-film probe used to measure u_{crit} . The flanges surrounding both the core and the sensor are mounted on the underside of the Perspex flume bed. *b*, Photomicrograph (Nomarski interference contrast, NIC) from surface sediments showing purple sulphur bacteria clumps attached to a sand grain. Scale bar, 50 μm . *c*, Photomicrograph (phase contrast) of surface sand grains connected by amorphous aggregates (*a*) associated with clumps of purple sulphur bacteria (*c*). Scale bar, 200 μm . *d*, Filamentous cyanobacteria (NIC) lying across sand grains from subsurface blue-green sediment horizon. Scale bar, 100 μm . Clumps of purple sulphur bacteria (*c*) are also visible. *e*, Sediment core (7.5-cm diameter) from a mat of purple sulphur bacteria after erosion in the flume. Flow direction was from right to left. The adjacent line drawing of the same core (*f*) indicates vertical lamination of the remaining undisturbed purple surface (*p*), underlying blue-green layer (*g*), and black reducing sediments (*b*) beneath (*g*).

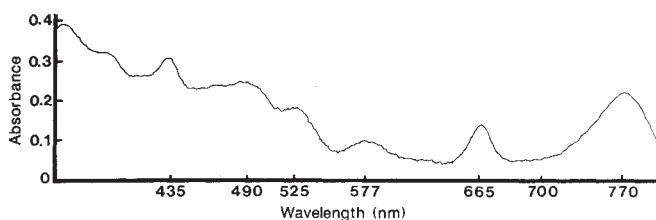


Fig. 2 Representative absorbance spectrum of acetone-extracted surface sediment sample (upper 2 mm layer) from a mat of purple sulphur bacteria, Pinellas Point, Florida.

(Fig. 1c), increasing grain cohesion. These bacteria may be *Thiocystis*, one of the few genera forming cell clumps and excreting slime⁹. Green subsurface sediments contained abundant motile cyanobacteria such as *Spirulina* and *Oscillatoria*, as well as clumps of purple bacteria (Fig. 1d). Polysaccharide sheaths and matrices produced by cyanobacterial trichomes¹¹ are known to contribute to sediment binding¹².

To quantify sediment binding by microbial mats, intact sediment cores (7.5-cm diameter, 10–12 cm sand, 10–11 cm supernatant water) collected at high tide were eroded in a 13-m-long recirculating laboratory flume with two-dimensional, fully developed, smooth turbulent flow. Sediment cores with obvious meiofauna or macrofauna bioturbation were discarded. Cores were inserted through the bottom of the flume into standing natural seawater (7-cm depth, 20–21 °C, 23–24 ‰), then pushed up through the core liner until the sediment surface was flush with the flume bed. Flow was slowly increased and u_{*crit} measured with a flush-mounted hot-film probe^{13,14} (maximum error 5%). Two types of erosion thresholds were observed. For minor microbial binding, for instance in the clear sediment cores, u_{*crit} was defined as the movement of 10 or more mucus-sand aggregates (see also below) similar to the definition for erosion of sterile sand in control cores (movement of 10 or more sand grains). In sediment cores with well-developed mats of purple or cyanobacteria, no single-grain movements occurred; instead, large pieces of laminae and embedded sand were torn up in carpet-like fashion (Fig. 1e, f). After the erosion threshold was reached, flow was halted and remaining undisturbed surface and subsurface sediment was sampled for chemical analyses.

We first examined pigment content of the microbial mat as a biochemical parameter to compare with the wide range of u_{*crit} observed in flume cores. The relation between photosynthetic pigments and cellular biomass is not constant owing to changing light and other conditions, but for specific microbial groups, pigments such as bacteriochlorophyll provide a quantitative index of microbial biomass¹⁵. Pigments were extracted from sediment samples in 90% acetone and analysed spectrophotometrically^{6,16}. Absorbance spectra from sediment samples revealed a variety of distinct peaks representing bacteriochlorophyll *a* (BChl *a*, A_{770}), chlorophyll *a* (Chl *a*, A_{665}), and other peaks identifiable only with chromatography (Fig. 2).

The value of u_{*crit} of bio-stabilized sediment ($n=21$) was correlated with the pigment concentration of purple sulphur bacteria (BChl *a*) in the top exposed 2 mm ($r=0.80$, $p<0.001$, Fig. 3). Chlorophyll *a* also showed a significant correlation with u_{*crit} ($r=0.79$, $p<0.001$). Chlorophyll *c* as an indicator of diatoms and their potential binding effects^{5,6} was not correlated with u_{*crit} ($r=-0.06$). We have emphasized the relation between BChl *a* and u_{*crit} (Fig. 3) because the BChl *a* peak is easily identified, and purple sulphur bacteria are predominant on the sediment surface at our sample site. BChl *a* is used only a first-order tool to provide a predictive correlate of biotic sediment stabilization.

Examination of the types of cores eroded shows that even in 'clear' surface sands, u_{*crit} was ~25% higher compared to sterile sediment (Table 1, columns 2 and 4). Underlying, previously subsurface, blue-green bacteria showed substantial stabilization

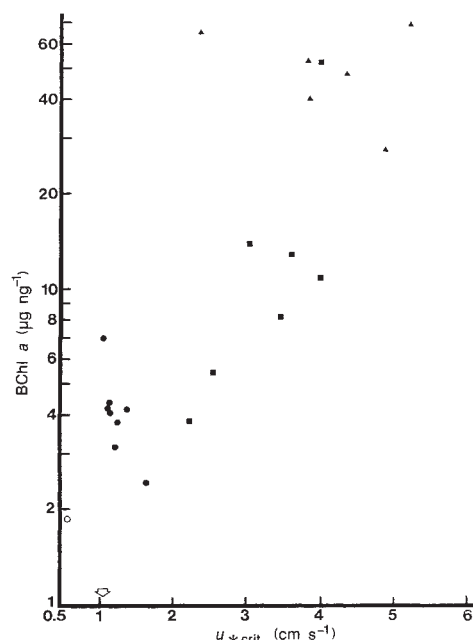


Fig. 3 The relation between sediment bacteriochlorophyll *a* in the upper 2 mm of exposed sediment and the critical friction velocity of biotically stabilized sediments in intact cores eroded in a laboratory flume. Microbial mats contain both purple sulphur bacteria and cyanobacteria and include cores as detailed in Table 1. The regression equation ($n=21$) is $\log Y = 0.29 X + 0.27$. Circles are 'clear' surface sediments; squares are blue-greens originally beneath the surface in the field but exposed at the surface in the flume by eroding overlying clear sand; triangles are purple surface sediments. The arrow on the abscissa indicates the mean u_{*crit} for sterile sediment.

when exposed by eroding the clear surface sands (Table 1, column 3). Cyanobacterial stabilization was comparable with that found on purple surfaces (column 1). Both purple sulphur bacteria and cyanobacteria, although morphologically different (Fig. 1), apparently contribute to sediment binding.

We stress that although photopigments are one approach to characterizing the mat, pigments do not cause the binding of sediment; they are indicators of microbial biomass which produces extracellular polymers. The relation between cellular biomass and the production of extracellular material is poorly known, especially in sediment bacteria. The close spatial association of purple and cyanobacteria¹⁷ (Fig. 1d), the correlation between BChl *a* and Chl *a* ($r=0.69$, $p=0.001$, $n=21$), and concurrent production of extracellular polymers by both groups^{10–12} all contribute to the significant relation between pigment concentrations and u_{*crit} .

Table 1 Critical friction velocity (u_{*crit}) of sediment cores (Pinellas Point) used in flume erosion experiments

	Natural cores			Control core
	Purple sulphur bacteria surface	'Clear' surface	Sediment subsurface (blue-greens)	Sterile surface
u_{*crit} (cm s ⁻¹)	4.06	1.28	3.66	1.04
SD	1.00	0.22	0.39	0.08
<i>n</i>	6	12	6	4

Purple (mats) and 'clear' (no visible mats) sediments were eroded from the surface of intact cores. Subsurface blue-green (cyanobacteria) laminae were eroded after erosive removal of overlying clear sediments. Sterilized control cores consisted of beach sediment washed and soaked in chlorine bleach.

Although there is no single measure that accounts for all extracellular biomass, many of these polymers include a polysaccharide component¹⁸. By agitating sediment samples, centrifuging to exclude cells and detritus, and analysing the supernatant (phenol-sulphuric acid hydrolysis), we have previously found elevated levels of 'colloid' carbohydrates corresponding to bacterial and microalgal mats^{5,6}. In this study, colloid concentrations (138–319 μg glucose equivalents per g sediment) were an order of magnitude greater than most values determined for a benthic diatom mat^{5,6}. At Pinellas Point, the carbohydrate index of extracellular material ($n=21$) was significantly correlated with Bchl a ($r=0.46$, $p=0.025$) and Chl a ($r=0.73$, $p=0.001$) as well as u_{crit} ($r=0.77$, $p<0.001$). These relations suggest that, in our experiments, pigment biomass indicators were related to the mass of extracellular material (carbohydrates). More direct assays of mat extracellular material are desirable, but are not readily available and may be too complex for practical engineering purposes. However, where bacterial mats are dominated by obvious groups of specific photoautotrophs, as commonly occurs with purple sulphur bacteria and cyanobacteria^{8,17}, pigment measures may provide useful predictors of u_{crit} . In mats where non-photosynthetic bacteria dominate (for example *Beggiatoa*), indicators such as sulphur content¹⁹ or colloid carbohydrate could be related to erosion criteria. We conclude that there are a variety of biological parameters that may form the basis of a predictive relation with critical erosion thresholds. This approach will provide more realistic threshold values for natural sediments of a given grain size than values obtained from criteria developed for abiotic sediments (for example Shields' curve; see Miller *et al.*¹).

To predict the stability and transport of natural sediments, determination of u_{crit} must be supplemented by information on the transport rates of eroded sediment. The extracellular material of well-developed mats had an armouring effect⁴, such that erosion occurred by a 'blow-out' phenomenon for $u_* > u_{\text{crit}}$, and the mat surface rapidly broke up into pieces that were transported as large aggregates. Even with less developed bacterial mats (clear cores), small aggregates of sand and mucus rather than individual grains were transported as bed-load. The flow eroded the underlying sediment, now exposed, until reaching laminae with a higher u_{crit} (for example cyanobacteria below clear sediment; Table 1). In these cases standard bed-load equations do not apply.

We obtained preliminary results on microbial binding and sediment transport rate from the mass of sediment collected in miniature traps (flush-mounted scintillation vials) downstream of cores. Bed-load transport from a control core with sterile sand ($u_* = 1.24 \text{ cm s}^{-1} > u_{\text{crit}}$) occurred at a rate of $3 \times 10^{-3} \text{ g cm}^{-1} \text{ s}^{-1}$. In contrast, three natural cores with clear surface sand were eroded at rates of $3\text{--}8 \times 10^{-4} \text{ g cm}^{-1} \text{ s}^{-1}$ at $1.20 < u_* < 1.65 \text{ cm s}^{-1}$. These initial experiments indicate that despite low mat development, sediment transport rate can be reduced by an order of magnitude compared with sterile sand transport at a similar or even smaller u_* . Detailed studies of organismal effects on transport rates are lacking, and the reader is cautioned that adjusting u_{crit} is only the initial step in accounting for biotic factors in models of natural sediment transport.

The laminated microbial community we have studied (Fig. 1F; 'Farbstreifensandwatt' or 'versicoloured sandy tidal flat'¹⁷) is common in shallow tropical and temperate sandy sediments⁸. Extensive mats of sulphur bacteria are also known from continental shelves and the deep sea^{8,20}. The effects reported here on the erosion and transport of biotically stabilized sediments will be widespread, with major implications for sedimentology, seabed engineering and disposal, nutrient flux¹⁹ and numerous other studies.

We thank Kelly Walker for assistance with field and laboratory experiments, Walter Bowles for grain-size analysis, John Paul for advice on microscopy, Bob Byrne for the use of scanning spectrophotometers, Ferenc Stefani for help with the photo-

graphic plate, and Craig Emerson and Jim Novitsky for valuable comments on the manuscript. This work was supported by the Natural Sciences and Engineering Research Council of Canada.

Received 29 April; accepted 6 September 1987.

1. Miller, M. C., McCave, I. N. & Komar, P. D. *Sedimentology* **24**, 507–527 (1977).
2. Dyer, K. R. *Coastal and Estuarine Sediment Dynamics* (Wiley, Chichester, 1986).
3. McCave, I. N. *The Benthic Boundary Layer* (Plenum, New York, 1976).
4. Jumars, P. A. & Nowell, A. R. M. *Cont. Shelf Res.* **3**, 115–130 (1984).
5. Grant, J. in *Tide Influenced Sedimentary Environments and Facies* (eds de Boer, P. J., van Gelder, A. & Nio, S. D.) (Reidel, The Netherlands, in the press).
6. Grant, J., Bathmann, U. V. & Mills, E. L. *Estuar. coast. Shelf Sci.* **23**, 225–238 (1986).
7. Rhoads, D. C. & Boyer, L. F. in *Animal–Sediment Relations* (eds McCall, P. L. & Tevesz, M. J. S.) **3**, 52 (Plenum, New York, 1982).
8. Cohen, Y., Castenholz, R. W. & Halvorson, H. O. *Microbial Mats, Stromatolites* (Liss, New York, 1984).
9. Trüper, H. G. & Pfennig, N. in *The Prokaryotes* (eds Starr, M. P., Stolp, H., Truper, H. G., Balows, A. & Schlegel, H. G.) **229–312** (Springer, Berlin, 1981).
10. Hirsch, P. in *Microbial Adhesion and Aggregation* (ed. Marshall, K. C.) **373–394** (Springer, Berlin, 1984).
11. Jorgensen, B. B., Revsbech, N. P. & Cohen, Y. *Limnol. Oceanogr.* **28**, 1075–1093 (1983).
12. Neumann, A. C., Gebelein, C. D. & Scoffin, T. P. *J. sedim. Petrol.* **40**, 274–297 (1970).
13. Gust, G. & Weatherly, G. L. *J. geophys. Res.* **90**, 4779–4792 (1985).
14. Kreplin, H. P. & Eckelmann, H. *Phys. Fluids* **22**, 1233–1239 (1979).
15. van Gemerden, H. *Arch. Mikrobiol.* **64**, 103–110 (1968).
16. Takahashi, M. & Ichimura, S. *Limnol. Oceanogr.* **15**, 929–944 (1970).
17. Stal, L. J., van Gemerden, H. & Krumbein, W. E. *FEMS Microbiol. Ecol.* **31**, 111–125 (1985).
18. Characklis, W. G. & Cooksey, K. E. *Adv. appl. Microbiol.* **29**, 93–138 (1983).
19. Grant, J. & Bathmann, U. V. *Science* **236**, 1472–1474 (1987).
20. Gallardo, V. A. *Nature* **268**, 331–332 (1977).

Radiocarbon in dissolved organic matter in the central North Pacific Ocean

Peter M. Williams & Ellen R. M. Druffel*

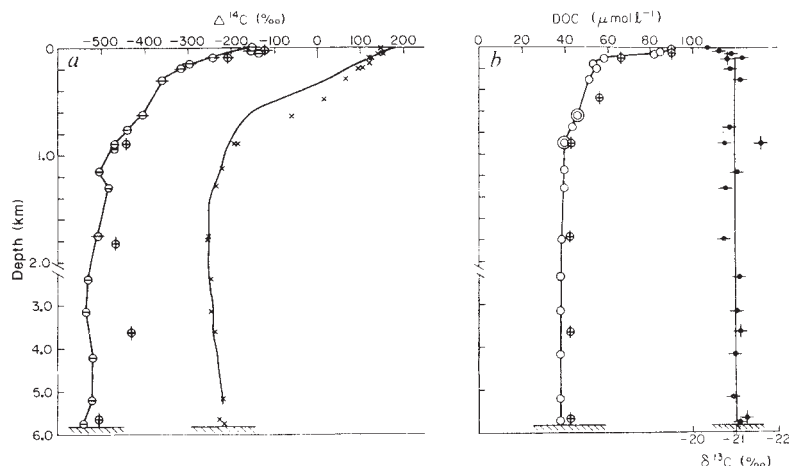
Scripps Institution of Oceanography, La Jolla, California 92093, USA

* Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA

The origin of dissolved organic carbon (DOC) in the ocean has been long debated. Whereas Mantoura and Woodward¹ have used the conservative nature of DOC in a British estuary to conclude that $\geq 50\%$ of DOC in the oceans could be river-derived, recent lignin results in the equatorial Pacific² have indicated that $\leq 10\%$ of the DOC is potentially of terrestrial origin. In addition, the $\delta^{13}\text{C}$ signature (relative to the PDB standard) of DOC ranges from -20 to -24% (refs 3, 4), indicating that the primary source of DOC is from marine-derived organic carbon. Here we present the first detailed profile of radiocarbon measured in DOC and dissolved inorganic carbon (DIC) in the oligotrophic gyre of the central North Pacific. $\Delta^{14}\text{C}$ (per mil deviation from the activity of 19th-century wood) of DOC ranged from -150% (1,310 yr BP) in surface waters to -540% (6,240 yr BP) at 5,710 m, 40 m off the bottom, where these 'apparent ages' or 'residence times' are mean values for the combined constituents of the DOC. The surprising similarity in the shapes of the profiles of $\Delta^{14}\text{C}$ in the DOC and DIC pools suggests that similar processes are controlling the radiocarbon distribution in each of the two reservoirs and that bomb-produced radiocarbon has penetrated the DOC + DIC pools to a depth of ~ 900 m. The depletion of the $\Delta^{14}\text{C}_{\text{DOC}}$ values by 300% with respect to the $\Delta^{14}\text{C}_{\text{DIC}}$ values suggests that a certain fraction of the DOC is recycled within the ocean on longer time-scales than DIC.

Radiocarbon in DOC and DIC was measured in 25 water samples collected from the central North Pacific Ocean (31°N , 159°W), one sample from each of 21 depths, with duplicate casts at 20, 100, 900 and 1,800 m. The water for DOC oxidation was filtered on board ship directly from 270-l stainless steel Gerard barrels through $1.0\text{-}\mu\text{m}$ pore-diameter, pre-combusted glass-fibre filters into cleaned glass bottles and frozen at -20°C . A 5-l sample was then thawed, acidified to pH 2.5 with H_3PO_4 , purged of DIC with pure N_2 gas, saturated with O_2 , and the DOC converted to CO_2 by photo-oxidation for 6 h

Fig. 1 *a*, $\Delta^{14}\text{C}_{\text{DOC}}$ ($\ominus$, $\oplus$) and $\Delta^{14}\text{C}_{\text{DIC}}$ ($\times$) values versus depth at 31°N , 159°W . The $\Delta^{14}\text{C}_{\text{DOC}}$ values were measured using TAMS techniques and are corrected to a $\delta^{13}\text{C}$ of -25‰ and are not yet corrected using the actual $\delta^{13}\text{C}$ values of the graphite. The error from counting statistics plus assumed graphite- $\delta^{13}\text{C}$ assignments and $\delta^{13}\text{C}$ values of the DOC samples range from ± 8 to $\pm 19\text{‰}$ and are indicated by horizontal error bars. Off-curve $\Delta^{14}\text{C}_{\text{DOC}}$ values ($\oplus$) are discussed in the text. The $\Delta^{14}\text{C}_{\text{DIC}}$ values, measured using conventional gas proportional counting techniques, are reported in the usual manner³⁰. The surface (3 m) value is the average of 12 surface $\Delta^{14}\text{C}_{\text{DIC}}$ values taken during the 35-day cruise in October 1985. The error from counting statistics is ± 2.9 to $\pm 3.3\text{‰}$. The smoothed curve represents $\Delta^{14}\text{C}_{\text{DIC}}$ results from GEOSECS Station 204, September 1973, located at $31^\circ 22' \text{N}$, $150^\circ 02' \text{W}$ (ref. 8). *b*, Total DOC concentrations ($\circ$, $\oplus$) and $\delta^{13}\text{C}_{\text{DOC}}$ values ($\bullet$, $\blacklozenge$) versus depth at 31°N , 159°W . DOC values are from manometric measurements of the CO_2 resulting from the photo-oxidation of DOC. The experimental error is $\pm 1.5 \mu\text{mol C l}^{-1}$. Double circles denote replicate samples from the same depth. Off-curve total DOC values ($\oplus$) are discussed in the text. $\delta^{13}\text{C}$ values were measured separately at the Woods Hole Oceanographic Institution. The experimental error, $\pm 0.15\text{‰}$, is given by the horizontal bars, where the vertical bars indicate values corresponding to the off-curve $\Delta^{14}\text{C}_{\text{DOC}}$ values. The vertical line is the mean value of all 20 points.



at 75°C in an all-glass reactor, according to published techniques^{3,5}. DIC was absorbed in $\text{SrCl}_2\text{-NH}_4\text{OH}$ on board ship from 200-l, unfiltered, acidified seawater samples⁶ for subsequent proportional counting at the Woods Hole Oceanographic Institution Radiocarbon Laboratory.

The $\Delta^{14}\text{C}_{\text{DOC}}$ measurements were made by A. J. T. Jull *et al.*⁷ by tandem accelerator mass spectrometry (TAMS). Profiles of 25 $\Delta^{14}\text{C}_{\text{DOC}}$ and 24 $\Delta^{14}\text{C}_{\text{DIC}}$ results are shown in Fig. 1*a*, along with $\Delta^{14}\text{C}_{\text{DIC}}$ values measured in 1973 during the GEOSECS Pacific expedition⁸. Profiles of total DOC concentrations and their corresponding $\delta^{13}\text{C}$ values are given in Fig. 1*b*. $\Delta^{14}\text{C}_{\text{DOC}}$ values range from -150‰ at 3 m to -540‰ at 5,710 m. These $\Delta^{14}\text{C}$ values are 100–200‰ lower than earlier results^{5,9} measured in the northeastern Pacific. We believe this discrepancy is a result of incomplete oxidation (recoveries were $\sim 70\%$) of DOC in the 500-l samples processed earlier.

The difference between $\Delta^{14}\text{C}_{\text{DOC}}$ in surface and deep waters (390‰) is exactly the same as that observed between surface (150‰) and deep (-240‰) $\Delta^{14}\text{C}_{\text{DIC}}$ from the same water samples, although the DOC values are 300‰ more negative (Fig. 1*a*). The very old apparent ages of DOC, coupled with constant $\delta^{13}\text{C}_{\text{DOC}}$ values (Fig. 1*b*), indicate that there is a recycled component of DOC which is resistant to oxidation. This is supported by the invariant concentration of DOC in deep waters of the North Pacific and Atlantic oceans (L. I. Gordon, personal communication and ref. 10), as shown in Fig. 1*b* for the central North Pacific, and its resistance to rapid bacterial utilization^{11,12}.

At present, it is possible to characterize only 20–30% of the organic matter in DOC. The sum of the total (free plus hydrolysable) soluble amino acids (measured fluorometrically¹³) and carbohydrates (measured spectrophotometrically¹⁴) at 31°N , 159°W is equivalent to $\sim 6\%$ of the total DOC below 1,000 m and 16% at the surface, and this may be considered the most labile (modern) fraction of the DOC. The higher $\delta^{13}\text{C}$ values in the upper 20 m (Fig. 1*b*) probably reflect the greater content of labile, modern constituents in surface waters. Humic material (fulvic and humic acids), which can be isolated by absorption onto XAD macroporous resins, constitutes 5–20% of the DOC in deep and surface waters of the open ocean^{2,15,16}, while lipids account for 2–5% of the DOC^{17,18}. The remaining 60–80% of the DOC has not yet been characterized, primarily due to the difficulties inherent in isolating polar organic material from dilute saline solutions.

The fact that the $\Delta^{14}\text{C}_{\text{DOC}}$ and $\Delta^{14}\text{C}_{\text{DIC}}$ curves (Fig. 1*a*) are almost identical in shape suggests that similar processes are controlling the radiocarbon distributions in each of these two reservoirs. It is commonly believed that the decrease in $\Delta^{14}\text{C}_{\text{DIC}}$

observed in North Pacific Deep Water represents the transit time of North Atlantic Deep Water from its origin in the North Atlantic to the North Pacific¹⁹, with consideration of the short circuit introduced by common water formation in the Antarctic²⁰. The similarity strongly suggests that a fraction of the DOC within the deep-water circulation is cycled in much the same way as DIC.

There are seven samples whose DOC concentrations (denoted by crossed circles, Fig. 1*b*) are 6–14% higher than the other corresponding DOC values. These higher concentrations were observed only in the first seven samples photo-oxidized. This cannot be due to increased oxidation of more refractory, older DOC as the $\Delta^{14}\text{C}$ values of six of these seven samples (Fig. 1*a*, crossed circles) were equal to or greater than the $\Delta^{14}\text{C}_{\text{DOC}}$ trend of the remaining samples and reoxidation of several of these surface and deep-water counterparts for an additional 6 h yielded 1% or less additional CO_2 . We suspect the higher DOC concentrations and $\Delta^{14}\text{C}$ values of these seven samples were due to modern, post-bomb organic matter ($\Delta^{14}\text{C} = 100\text{--}200\text{‰}$) that was photo-oxidized only in samples processed using the initial UV lamp. Subsequent oxidations with this same lamp or new lamps did not duplicate the earlier results. However, known compounds (for example, *N*-acetylglucosamine and serum albumin) were 100% oxidized with all lamps, regardless of their histories.

Intriguing deviations from the smooth curve are apparent in the $\Delta^{14}\text{C}_{\text{DOC}}$ results (Fig. 1*a*). There is a suggestion of a minimum in $\Delta^{14}\text{C}$ at 1,150 m, coincident with the lower portion of the oxygen minimum (950 m). The deepest sample, 40 m off the bottom, was significantly lower in $\Delta^{14}\text{C}$ than the other deep-water samples, suggesting that old DOC derived from low ^{14}C -activity sedimentary organic matter^{9,21} may be exchanging with DOC in the water column. The increase in $\Delta^{14}\text{C}_{\text{DOC}}$ of 100‰ at 3,600 m, if valid, may reflect increased sea-floor mineralization of modern particulate organic matter at this depth, as postulated by Jahnke and Jackson²². These deviations from the smooth $\Delta^{14}\text{C}_{\text{DOC}}$ trend are being verified by further measurements.

The average of four $\Delta^{14}\text{C}_{\text{DOC}}$ values in the upper mixed layer (~ 40 m) was -146‰ . This could result from a mixture of 44% old, deep-water DOC (average $\Delta^{14}\text{C} = -525\text{‰}$) and 56% post-bomb, surface DOC derived during the past 30 yr from planktonic biocarbon (average $\Delta^{14}\text{C} = 150\text{‰}$). In this model, of the $87 \mu\text{mol C l}^{-1}$ (average concentration of DOC in the surface mixed layer, Fig. 1*b*), $38 \mu\text{mol C l}^{-1}$ represents the recycled fraction of deep-water DOC resistant to rapid bacterial utilization, and $49 \mu\text{mol C l}^{-1}$ is the labile fraction continuously

utilized by microplankton and heterotrophic bacteria. The fact that the DOC content of oceanic surface waters is never less than $\sim 35 \mu\text{mol C l}^{-1}$ (ref. 23) can be attributed to the presence of this recycled, refractory DOC component.

The $\Delta^{14}\text{C}_{\text{DIC}}$ results agree well with a previous profile made during the GEOSECS study⁸ 12 years previously (Fig. 1a). Our data show a deeper penetration of the bomb radiocarbon transient to 900 m, ~ 100 –200 m deeper than that measured in 1973. However, there is essentially no difference between our $\Delta^{14}\text{C}$ values deeper than 1,100 m and comparable GEOSECS measurements. Similarity of the $\Delta^{14}\text{C}_{\text{DIC}}$ and $\Delta^{14}\text{C}_{\text{DOC}}$ curves strongly suggest that bomb radiocarbon has also penetrated the DOC pool to 900 m depth, most likely by advective and diffusive transport along isopycnal surfaces. An increase in $\Delta^{14}\text{C}_{\text{DIC}}$ at 3,600 m in concert with that observed in the $\Delta^{14}\text{C}_{\text{DOC}}$ profile is not expected, as the concentration of DIC is 50 times that of DOC in the deep sea, and thus much less susceptible to small changes.

Assuming a steady-state concentration and no terrestrial input of DOC, the apparent age of 6,000 yr for the DOC is also an indication of the amount of photosynthetically fixed carbon which enters the deep sea as DOC and is concurrently combusted to CO_2 . The average amount of carbon fixed photosynthetically in the world's open oceans (excluding adjacent seas) has been estimated at between $130 \text{ g C m}^{-2} \text{ yr}^{-1}$ (ref. 24) and $500 \text{ g C m}^{-2} \text{ yr}^{-1}$ (ref. 25), amounting to a global production of approximately 40 – $150 \times 10^{15} \text{ g yr}^{-1}$. At $40 \mu\text{mol l}^{-1}$ of DOC, the total reservoir of DOC in the deep sea (vol. $= 1.3 \times 10^{21} \text{ l}$ for 300–4,100 m depth²⁶) is about $6 \times 10^{17} \text{ g}$, and thus the steady-state input of DOC is $0.1 \times 10^{15} \text{ g C yr}^{-1}$ ($6 \times 10^{17} \text{ g C}$ per 6,000 yr) or about 0.067–0.25% of the photosynthetically fixed carbon. This input of DOC is 0.5–3% of the amount of particulate organic carbon falling out of the upper 300 m into the deep ocean^{24,25,27} (new production), assuming the above global carbon production rates.

If 50% of the DOC in the deep sea were derived from global riverine sources¹, then it would require $\sim 1,500$ yr to fill the ocean with respect to terrestrial DOC at a riverine flux of $2 \times 10^{14} \text{ g C yr}^{-1}$ (ref. 28), assuming that riverine DOC entering the ocean is retained as DOC and that it is in steady-state concentration. The average $\Delta^{14}\text{C}$ value for soluble humic and fulvic acids collected between 1982 and 1984 from the Amazon River system²⁹ is 270‰, indicating that terrestrially-derived DOC is modern in age. Thus a potential terrestrial DOC component, even during the pre-bomb period when $\Delta^{14}\text{C}$ was supposedly 0‰, would have to be recycled at least as many times as marine-derived DOC in the world oceans to give the resultant low $\Delta^{14}\text{C}_{\text{DOC}}$ values in the deep sea. We believe that the absence of significant amounts of lignin in open-ocean humic substances² coupled with the constraints imposed by marine $\Delta^{14}\text{C}$ and $\delta^{13}\text{C}$ values, precludes a terrestrially derived DOC component exceeding 10% of the total DOC.

In summary, we speculate that a substantial fraction of DOC, regardless of its origin, may be common to all oceanic deep waters and undergoes multiple recycling, similar to the recycling of salt in the ocean. The apparent mean ages or residence times of about 6,000 yr BP could result from mixing 51% dead DOC ($\Delta^{14}\text{C} = -1,000$ ‰) with 49% pre-bomb DOC (-40 ‰), neglecting any recent input of bomb radiocarbon. However, the striking similarity between the $\Delta^{14}\text{C}_{\text{DIC}}$ and $\Delta^{14}\text{C}_{\text{DOC}}$ curves suggests that there is indeed bomb radiocarbon in the DOC pool, at least in the upper 900 m of the central North Pacific Ocean. Our reoccupation of this site in June 1987 will help resolve this question, as further penetration of the bomb radiocarbon signal into the DOC pool is anticipated 18 months later.

We thank Drs A. Jull, D. Donahue and T. Linick for analysing DOC samples at the University of Arizona TAMS facility, S. Griffin and J. Frankenthal for measuring $\Delta^{14}\text{C}_{\text{DIC}}$, K. Robertson for analysing amino acids and carbohydrates, John Hedges and Rick Jahnke for helpful comments and R. Paddock, K. Smith

and the crew of the R V *Melville* for help with sample collection and ship logistics. Financial support was provided by the NSF. Woods Hole Oceanographic Institution Contribution No. 6593.

Received 12 June; accepted 11 September 1987.

1. Mantoura, R. F. C. & Woodward, E. M. S. *Geochim. cosmochim. Acta* **47**, 1293–1309 (1984).
2. Meyers-Schulte, K. J. & Hedges, J. I. *Nature* **321**, 61–63 (1986).
3. Williams, P. M. & Gordon, L. I. *Deep-Sea Res.* **17**, 19–27 (1970).
4. Eadie, B. J., Jeffrey, L. M. & Sackett, W. M. *Geochim. cosmochim. Acta* **42**, 1265–1269 (1978).
5. Williams, P. M., Oeschger, H. & Kinney, P. *Nature* **224**, 256–258 (1969).
6. Linick, T. W. *Radiocarbon* **28**, 559–607 (1980).
7. Jull, A. J. T., Donahue, D. J., Hathaway, A. L., Linick, T. W. & Toolin, L. J. *Radiocarbon* **28**, 191–197 (1986).
8. Ostlund, H. G. & Stuiver, M. *Radiocarbon* **22**, 25–40 (1980).
9. Williams, P. M., Stenhouse, M. C., Druffel, E. M. & Koide, M. *Nature* **265**, 698–701 (1978).
10. Williams, P. M., Carlucci, A. F. & Olson, R. *Oceanologica Acta* **3**, 471–476 (1980).
11. Barber, R. T. *Nature* **220**, 274–275 (1968).
12. Williams, P. M. & Carlucci, A. F. *Nature* **262**, 810–811 (1976).
13. Lindroth, P. & Mopper, K. *Analyt. Chem.* **51**, 1667–1674 (1979).
14. Burney, C. M. & Sieburth, J. McN. *Mar. Chem.* **5**, 15–28 (1977).
15. Stuermer, D. H. & Harvey, G. R. *Nature* **250**, 480–481 (1976).
16. Gagosian, R. B. & Stuermer, D. H. *Mar. Chem.* **5**, 605–632 (1977).
17. Williams, P. M. *J. Fish. Res. Bd Canada* **22**, 1107–1122 (1965).
18. Kennicutt II, M. C. & Jeffrey, L. M. *Mar. Chem.* **10**, 367–387 (1981).
19. Broecker, W. S. *Chemical Oceanography* (Harcourt, Brace, Jovanovich, New York, 1974).
20. Stuiver, M., Quay, P. D. & Ostlund, H. G. *Science* **219**, 849–854 (1983).
21. Druffel, E. M., Williams, P. M., Livingston, H. & Koide, M. *Earth planet. Sci. Lett.* **71**, 205–214 (1984).
22. Jahnke, R. A. & Jackson, G. A. *Nature* **329**, 621–623 (1987).
23. Jackson, G. A. & Williams, P. M. *Deep-Sea Res.* **31**, 223–235 (1985).
24. Martin, J. H., Knauer, G. A., Karl, D. M. & Broenkow, W. W. *Deep-Sea Res.* **34**, 267–287 (1987).
25. Jenkins, W. J. *Nature* **300**, 246–248 (1982).
26. Sverdrup, H. U., Johnson, M. W. & Fleming, R. H. *The Oceans* (Prentice-Hall, New York, 1942).
27. Pace, M. L., Knauer, G. A., Karl, D. M. & Martin, J. H. *Nature* **325**, 803–804 (1987).
28. Meybeck, M. *Am. J. Sci.* **282**, 401–450 (1982).
29. Hedges, J. I. *et al. Science* **231**, 1129–1131 (1986).
30. Stuiver, M. & Polach, H. *Radiocarbon* **19**, 355–363 (1977).
31. Craig, H., Broecker, W. S. & Spencer, D. *GEOSECS Pacific Expeditions Vol. 4* (NSF, Washington DC, 1981).

Is the periodicity of extinctions a taxonomic artefact?

Colin Patterson & Andrew B. Smith

Department of Palaeontology, British Museum (Natural History), London SW7 5BD, UK

The evidence that extinctions since the mid-Permian show periodicity of ~ 26 million years rests primarily on the stratigraphic ranges of marine families. We have checked the echinoderm and fish families which together make up about 20% of the data. Only 25% of these fish and echinoderm extinctions are real (disappearance of a monophyletic group). The remaining 75% is noise, chiefly 'extinctions' of non-monophyletic groups, mistaken dating, and 'families' containing one species only. The signal-to-noise ratio is very similar in echinoderms (27:73) and fishes (23:77). Periodicity in our sample is a feature of the noise component, not of the signal.

Periodicity of extinctions has received much attention recently, both as concerns the phenomenon^{1–18} and possible explanations for it^{12–26}. Raup and Sepkoski¹ initially analysed a subset culled from Sepkoski's *Compendium of Fossil Marine Families*²⁷ covering the last 250 million years (Myr) (late Permian to mid-Miocene). They omitted all extant families, all whose last occurrences were not known to stage, and all with questionable taxonomic or stratigraphic designation. In their opinion¹, this culling 'removed much of the noise that characterizes data sets of this kind'. The 567 families analysed comprised roughly 35% cephalopods, 11% vertebrates, 10% echinoderms, and 6–8% each bivalves, brachiopods, bryozoans, sponges and gastropods. All extinctions were treated as occurring at the end of the stage in which each taxon is last recorded. Twelve extinction peaks were found, and when fitted to the Harland *et al.*²⁸ timescale, a 26 Myr cycle centred on the end-Cretaceous extinction coincided with seven of the 11 other peaks. Statistical testing showed that this is highly significant.

Subsequently, Raup and Sepkoski^{2,3} analysed more extensive data comprising about 1,800 (ref. 2) and 2,160 (ref. 3) families. These increases were due to incorporation of roughly 1,000 extant families (which cannot affect inferred extinctions), about 150 families with extinction times known only to series or system (which cannot affect inferred extinctions at stage level), some 150 newly recognized or recorded extinct families, and over 200 extinctions in late Permian and post-mid-Miocene stages excluded earlier. Raup and Sepkoski have also analysed an unpublished compilation of marine invertebrate genera^{3,4}. Applying varied measures of extinction intensity and statistical tests, this further work generally confirms the 26 Myr periodicity, but only eight family extinction peaks are significantly greater than the background intensity², and these eight are those most closely matching the 26 Myr cycle. The data on genera^{3,4} indicate the same eight mass extinctions, but also suggest a Pliocene event (not predicted by the 26 Myr cycle), an Aptian event (predicted by the 26 Myr cycle), and a Carnian peak (attributed to sampling error).

Evaluation^{5,10,12,14-16} of the 26 Myr periodicity thesis, and responses to it⁷⁻⁹, have so far centred on the way the data were treated, and on the timescale chosen. Here we question a different aspect, the biological reality of the extinctions represented by the echinoderm and fish families used in Raup and Sepkoski's original sample of 567 Dzhulfian to mid-Miocene families¹ and their later^{2,3} analyses of 973 Leonardian to Pleistocene families (which comprised 25% cephalopods, 12% each echinoderms and vertebrates, and 7-9% each brachiopods, sponges, gastropods and bivalves). We do not address the generic data here, though we are preparing an analysis. Raup and Sepkoski^{3,4,18} acknowledge that their data contain noise, but argue that 'nonexistent extinction events will degrade the sample...towards randomness'³ and 'uncertain data can destroy a regular pattern but they cannot create one'¹⁸. As we shall show, this supposition that noise is random is not always met.

Sepkoski's *Compendium* necessarily relies primarily on secondhand data²⁹⁻³¹. Every specialist finds faults in such compilations, and there have been important changes recently in our view of taxa such as the extinct families listed in the *Compendium*. These changes centre on the acceptance of cladistic method and terminology, where for our topic the significant change is in our understanding of monophyletic groups (which can become extinct) and non-monophyletic groups (which cannot). The essence of the distinction between the two is that monophyletic groups comprise real genealogical entities forming individual branches of the phylogenetic tree, whereas non-monophyletic groups are created by taxonomists, and as their limits are artificially imposed their 'disappearance' does not directly reflect extinction of a lineage.

Sepkoski and Raup² commented that 'families are rather large and arbitrary taxonomic units, that on average contain about

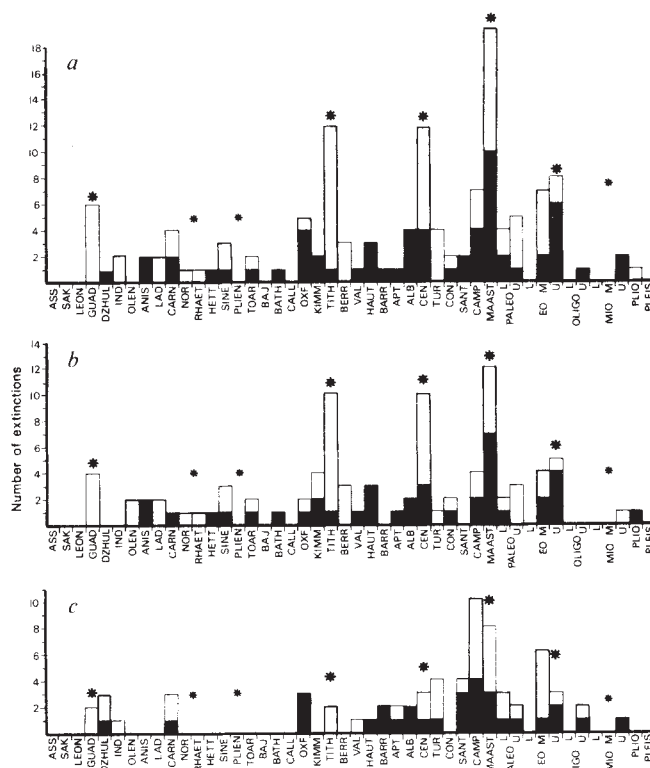


Fig. 1 Plot of the number of extinctions of echinoderm (filled bars) and fish (open bars) families in each stage from the Guadalupian (Upper Permian, crinoids not plotted) to Recent. *a*, Data taken from Sepkoski²⁷ and 1984-1986 supplements ($n = 132$). *b*, Data from the same source that we reject as noise ($n = 91$). *c*, Data from the same source that we accept as representing valid extinctions (including corrected data where the last occurrence was wrongly attributed and monophyletic elements within paraphyletic and polyphyletic families) ($n = 68$). The eight extinction peaks identified by Sepkoski and Raup² are asterisked; large asterisks indicate those peaks that are also conspicuous in our sample, small asterisks those that are not, and which must be based on other taxonomic groups.

10^2 species'. Such units may become 'extinct' in only two ways, either by the disappearance of a distinctive, monophyletic group of species or by imposition of an arbitrary cut-off point creating a non-monophyletic group, a taxonomist's artefact.

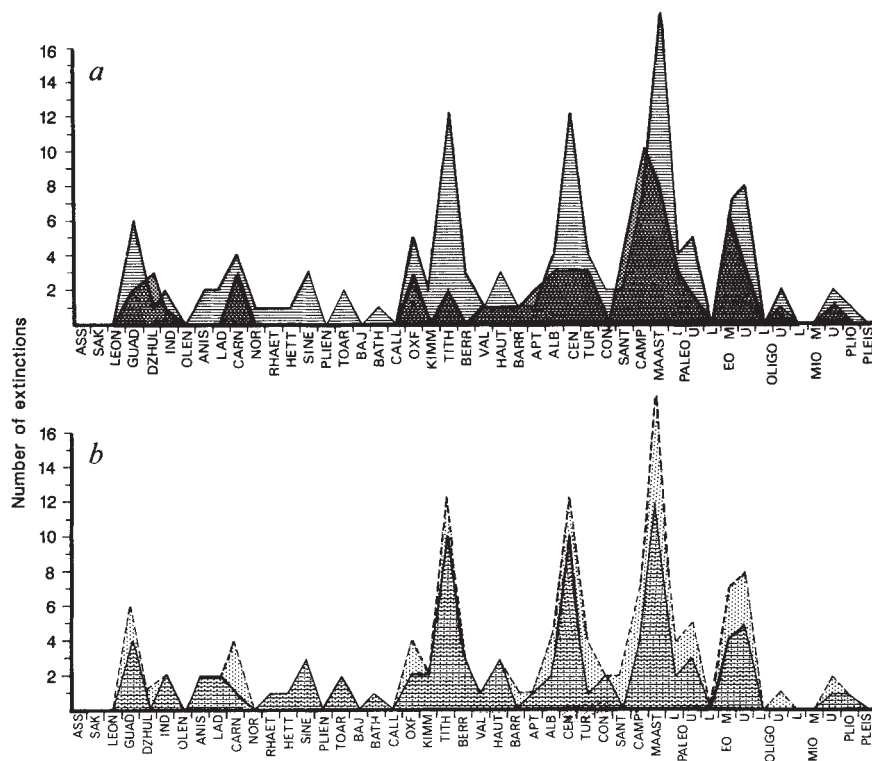
As specialists on echinoderms and on fishes, we knew few extinct families contain as many as 10^2 species, and believed that artificial boundaries to taxa abound. In our experience with fossil families, monophyly (and so the possibility of extinction) seemed to be almost the exception, not the rule. If so, the family

Table 1 Echinoderm and fish family extinctions in Sepkoski's *Compendium of Fossil Marine Families* and 1984-86 supplements

(<i>n</i>)	Echinoids (38)	Crinoids (25)	All echinoderms (66)	Chondrichthyes (19)	Osteichthyes (59)	All fishes (78)	Total (144)
Monophyletic, correct stage	26	28	27	26	22	23	25
Monophyletic, wrong stage	13	12	12	26	15	18	15
Paraphyletic	32	36	32	—	15	12	20
Polyphyletic	10.5	—	9	11	5	6	8
Non-monophyletic	5	20	10.5	11	8	9	10
Monotypic	13	4	9	26	24	24	17
Non-marine	—	—	—	—	10	8	4

144 families, 66 echinoderm and 78 fish, were checked, as follows: all families extinct during and since the late Permian in Echinoidea (38 families), Asteroidea (3 families), Chondrichthyes (19 families) and Osteichthyes (59 families), and in Crinoidea the 25 families extinct since the late Permian (we do not have the expertise necessary to check the numerous late Permian extinctions). The results are tabulated as %. Since 1984, J. J. Sepkoski has produced annual, cumulative *Corrections and Additions* to the compendium²⁷, available from the author.

Fig. 2 *a*, Plot of genuine extinctions in echinoderm and fish families in each stage, as recognized in this paper (stipple), superimposed on the total subset data taken from Sepkoski²⁷ and 1984–1986 supplements (hatched). *b*, Plot of noise in each stage, as recognized in this paper (hatched), superimposed on the total subset data as above (stipple).



extinction data contain much more noise than Raup and Sepkoski suspected. We therefore checked the data for echinoderms and fishes (Table 1, Fig. 1), 144 families in total. About 110 of these (20% of 567) would be included in Raup and Sepkoski's first analysis¹, and all were included in the later analyses^{2,3}, where they comprise 18% of about 800 extinctions dated to stage. The details of our analysis will be published elsewhere, and the results are summarized in Figs 1–3 and Table 1. We recognize seven categories of taxa in Raup and Sepkoski's data, as follows: (1) Monophyletic families containing two or more species whose last appearance is attributed to the correct stage. We accept these as valid extinctions. (2) Monophyletic families whose last appearance is attributed to the wrong stage because species wrongly attributed to the group extend the range, or because species are wrongly excluded, or through compiler's error. (3) Paraphyletic families, whose termination is what some workers^{32–34} speak of as pseudoextinction. Paraphyletic taxa are created when systematists hive off a derived portion of a monophyletic group, leaving the remainder as a paraphyletic 'ancestral group', united only by the possession of primitive features. Paraphyletic groups do not become extinct biologically, but by systematists' judgements. Van Valen³³ regards pseudoextinctions of this sort at family level as negligible, because some part of the ancestral group will persist after the birth of its descendant. Where we could recognize this, we treat it as valid extinction, but contrary to Van Valen's expectation these valid paraphyletic extinctions are negligible; we found only three among which two were incorrectly dated. Some regard paraphyletic groups as valid in extinction analysis for other reasons, either because they represent ('ecologically unified') groups, or because their last records provide a random sample of species extinctions. The first argument implies that extinction of niches rather than taxa is the phenomenon under analysis; the second could be tested by matching a reasonable sample of paraphyletic 'pseudoextinctions' against a comparable sample of randomly selected species extinctions. No proper sample of pseudoextinctions is yet available, but we predict that such a sample, like our sample of monotypic families (category 6 below), will mirror the quality of the fossil record, rather than genuine reductions

in diversity. (4) Polyphyletic families, including species which are not immediately related. Broadly, a polyphyletic group is one united by convergent rather than homologous features. As artificial groupings these obviously cannot become extinct. (5) Non-monophyletic families. These are not demonstrably monophyletic, and may be either para- or polyphyletic. In practice, discrimination between those two demands knowledge of phylogeny or of the taxonomist's intentions³⁵ which is often unavailable. (6) Monotypic families, containing only a single species. Here the family is an empty category used not to group species, but to express an opinion (usually concerning phenetic distinctness or ignorance of relationships) about the one included. In echinoderms and fishes monospecific families are generally members of clades with very poor fossil records, and the majority are known from only one horizon and locality. For such taxa it is of course possible that the unique occurrence and the extinction are coincident, but the weakness of that assumption is obvious, and we see these 'families' as representing exceptional fossilisation, not extinction. (7) Non-marine families. Some of the fish families are known only from non-marine deposits.

Our total fish and echinoderm data (Fig. 1*a*) show obvious peaks corresponding to five of the eight identified by Raup and Sepkoski: in the Guadalupian, Tithonian, Cenomanian, Maastriichtian and late Eocene. Fishes and echinoderms do not contribute to the peaks in the late Triassic, Pliensbachian, or Middle Miocene. The fish and echinoderm data may be divided into 'clade' and 'noise' components, where 'clade' means monophyletic groups with correct last records, category (1) above, and 'noise' includes families in categories (2)–(7). As Table 1 shows, the 'clade' component averages only 25%, and is similar in echinoderms (27%) and fishes (23%).

The 'noise' component of the data, though similar in total in fishes and echinoderms, has different sources in the two (Table 1). In echinoderms it comes mainly from paraphyletic ('ancestral') groups, but in fishes the dominant noise component is monotypic families. Mistaken dating of the last records of monophyletic groups makes a sizeable contribution in both echinoderms and fishes.

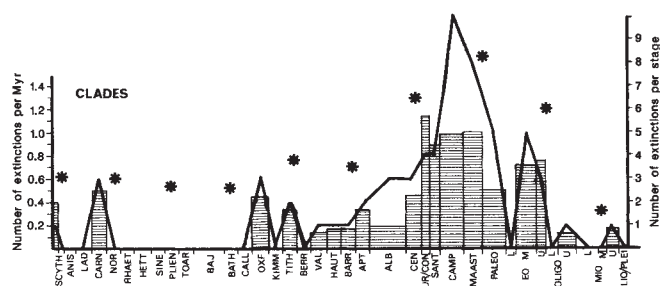


Fig. 3 Plot of genuine extinctions of echinoderms and fish families, as accepted in this paper, against the Harland *et al.*²⁸ timescale, stages from basal Triassic onwards. Continuous line, number of extinctions per stage; bar chart, number of extinctions per million years calculated for each stage. Asterisks indicate expected peaks of a 26 Myr cycle centered on the Cretaceous/Tertiary boundary; the end-Permian asterisk is displaced from the ordinate.

When plotted separately as 'clade' and 'noise' components (Figs 1*b, c* and Fig. 2), it is very obviously the noise that tracks the peaks in the total data, not the clades. The peaks evident in the total data are shifted in the clade plots: the Guadalupian, Tithonian and Cenomanian peaks are no longer evident, the Maastrichtian peak is lower than its neighbour in the Campanian, and the Upper Eocene peak is lower than that in the Middle Eocene. When the clades alone are plotted on the Harland *et al.*²⁸ timescale with Raup and Sepkoski's 26 Myr cycle superimposed (Fig. 3), there is no periodicity; instead, a late Cretaceous plateau is implied. That these results are not a fortuitous product of reduced sample size is indicated by the very different patterns of the 'noise' and 'clade' data, although these are comparable samples (respectively 11% and 9% of about 800 extinctions to stage in Sepkoski and Raup's analyses^{2,3}).

Our results indicate that (1), the fish and echinoderm families analysed by Raup and Sepkoski represent about 75% 'noise' (wrongly dated, or non-monophyletic, or one-species families) and only 25% 'signal' (correctly dated monophyletic groups), (2), those families show five of the eight extinction peaks identified by Raup and Sepkoski, but it is the noise, not the signal, which peaks in these five stages and (3), the clade data (Fig. 3) do not show periodicity. It follows that the periodicity apparent in Fig. 1*a* is not a real phenomenon in need of a biological explanation, but an artefact of taxonomic practice.

We believe that our taxonomists' test of the periodicity thesis raises two outstanding questions. First, whether our sample of Raup and Sepkoski's¹⁻³ data (~20% of their family extinctions to stage) is representative of the whole; and second, why the obvious peaks in our sample corresponding to five of the eight 'cyclic' peaks should be mostly noise—why should noise show periodicity? We believe that the first question can be answered in two ways. In terms of sampling it could be demonstrated statistically that our subset shows (or does not show) the same periodicity as the remainder of the sample; and in terms of data it could be shown that other groups contain (or do not contain) such a high proportion of noise. The latter is a job for specialists in other groups.

The second question, why noise should show periodicity, does not appear to have a simple or single answer as the noise in our data takes quite a different form in fishes and in echinoderms. We offer no clear answer to the question, but can speculate. Periodicity is evident in Sepkoski's generic data⁴, as well as in species data for echinoderms (ref. 36 and our own analysis), indicating that sampled species diversity fluctuates regularly. This could be a taphonomic feature, or a real pattern of extinction. The preponderance of false extinction in higher taxa (families, genera) after each peak indicates to us that taphonomy

and not extinction is the cause. Periodicity was first recognized on the basis of peaks in diversity³⁷⁻³⁹, and detailed analysis of our echinoderm and fish data (to be published elsewhere) implies that the noise component peaks when stages with a good fossil record are succeeded by stages with a poor one. If so, the problem of periodicity becomes taphonomic. Periodicity could conceivably also have a 'monographic' source, if specialists are reluctant to extend taxa across boundaries first established at European lithofacies changes. As Thomson⁴¹ found with fish diversity patterns, 'paleoichthyologists have been unwilling to admit that any genus could cross a period boundary'.

We are grateful to David Raup and Jack Sepkoski for unpublished data, and to Jack Sepkoski for constructive criticism.

Received 15 April; accepted 17 September 1987.

1. Raup, D. M. & Sepkoski, J. J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 801-805 (1984).
2. Sepkoski, J. J. & Raup, D. M. in *Dynamics of Extinction* (ed. Elliott, D. K.) 3-36 (Wiley, New York, 1986).
3. Raup, D. M. & Sepkoski, J. J. *Science* **231**, 833-836 (1986).
4. Sepkoski, J. J. in *Patterns and Processes in the History of Life* (eds Raup, D. M. & Jablonski, D.) 277-295 (Springer, Berlin, 1986).
5. Hoffman, A. & Ghiold, J. *Geol. Mag.* **122**, 1-4 (1985).
6. Hoffman, A. *Nature* **315**, 659-662 (1985).
7. Sepkoski, J. J. & Raup, D. M. *Nature* **321**, 533 (1986).
8. Gilinsky, N. L. *Nature* **321**, 533-534 (1986).
9. Kitchell, J. A. & Estabrook, G. *Nature* **321**, 534-535 (1986).
10. Hoffman, A. *Nature* **321**, 535-536 (1986).
11. Jablonski, D. *Paleobiology* **10**, 139-145 (1984).
12. Hallam, A. *Nature* **308**, 686-687 (1984).
13. Kerr, R. A. *Science* **227**, 1451-1453 (1985).
14. Kitchell, J. A. & Pena, D. *Science* **226**, 689-672 (1984).
15. Hoffman, A. N. *Jb. Geol. Paläont. Abh.* **172**, 219-244 (1986).
16. Connor, E. F. in *Patterns and Processes in the History of Life* (eds Raup, D. M. & Jablonski, D.) 119-147 (Springer, Berlin, 1986).
17. Flessa, K. W. *et al.* in *Patterns and Processes in the History of Life* (eds Raup, D. M. & Jablonski, D.) 235-257 (Springer, Berlin, 1986).
18. Raup, D. M. *Palaeontology* **30**, 1-13 (1987).
19. Benton, M. J. *Nature* **314**, 496-497 (1985).
20. Raup, D. M. *Science* **231**, 1528-1533 (1986).
21. Rampino, M. R. & Stothers, R. B. *Nature* **308**, 709-712 (1984).
22. Schwartz, R. D. & James, P. B. *Nature* **308**, 712-713 (1984).
23. Whitmire, D. P. & Jackson, A. A. *Nature* **308**, 713-715 (1984).
24. Davis, M., Hut, P. & Muller, R. A. *Nature* **308**, 715-717 (1984).
25. Alvarez, W. & Muller, R. A. *Nature* **308**, 718-720 (1984).
26. Whitmire, D. P. & Matese, J. J. *Nature* **313**, 36-38 (1985).
27. Sepkoski, J. J. *Contr. Biol. Geol. Milwaukee publ. Mus.* **51**, 1-125 (1982).
28. Harland, W. G. *et al.* *A Geological Time Scale* (Cambridge University Press, 1982).
29. Moore, R. C., Teichert, C. & Robison, R. A. (eds) *Treatise on Invertebrate Paleontology* (Geological Society of America and University of Kansas, Lawrence, Kansas, 1953-1986).
30. Harland, W. B. *et al.* (eds) *The Fossil Record* (Geological Society of London, 1967).
31. Romer, A. S. *Vertebrate Paleontology*, 3rd edn (University of Chicago Press, 1966).
32. Simpson, G. G. *The Major Features of Evolution* (Columbia University Press, New York, 1953); *Principles of Animal Taxonomy* (Columbia University Press, New York, 1961).
33. Van Valen, L. M. *Evol. Theory* **1**, 1-30 (1973); *Nature* **307**, 50-52 (1984).
34. Stanley, S. M. *Macroevolution: Pattern and Process* (Freeman, San Francisco, 1979).
35. Hennig, W. *Syst. Zool.* **24**, 244-256 (1975).
36. Kier, P. M. *J. Paleont.* **48** (Suppl. to No. 3: Mem. 5), 1-95 (1974).
37. Fischer, A. G. & Arthur, M. A. *Spec. Publ. Soc. econ. Paleont. Miner. Tulsa* **25**, 19-50 (1977).
38. Thomson, K. S. *Nature* **261**, 578-580 (1976).
39. Thomson, K. S. *Deus Palaeont. Stratigr.* **5**, 377-404 (1977).

SEPKOSKI REPLIES—Patterson and Smith¹ have performed a very valuable service in evaluating the palaeontological data on families of post-Palaeozoic echinoderms and vertebrates that are assembled in my *Compendium of Fossil Marine Families*². I greatly appreciate the errors and imprecisions they have brought to my attention and the general comments they have made on the taxonomic styles incorporated into currently available classifications. I hope these initial efforts will be followed by more detailed taxonomic treatments formalizing their present opinions.

I find fault, however, with their conclusions concerning the use of familial data for discerning extinction patterns and their implication that periodicity of extinction may be an artefact of flawed taxonomy. They judge that only ~25% of traditionally recognized families of post-Palaeozoic echinoderms and vertebrates represent monophyletic clades with two or more species. These they consider to be real data (or 'signal') capable of reflecting true extinction patterns. They judge the remainder of the data on post-Palaeozoic echinoderms and vertebrates to be 'noise', consisting of misdated, non-monophyletic, and/or monotypic families. When such families are excluded and only

polytypic, strictly monophyletic families are analysed, they find no evidence of a 26 Myr periodicity in extinction. They conclude that any periodicity evident in echinoderm and vertebrate families is an artefact of taxonomy and not a biological phenomenon.

At the core of Patterson and Smith's argument seems to be an assumption that taxonomic families are being treated as real evolutionary entities in the analyses of extinction patterns and that families can be subject to biological processes of extinction. This notion is mistaken. Extinction, as a dynamical process, operates at the population and (bio)species level. The higher taxa recognized by taxonomists become extinct only when all of their constituent species happen to be extinguished.

Throughout my analyses of taxonomic data, I have treated families as convenient proxies for species, either implicitly^{3,4} or explicitly⁵⁻¹¹. Each familial extinction, if properly dated, represents the extinction of one or more species; a collection of families therefore provides a sample of species extinctions, as Patterson and Smith note. The evidence that families provide a good sample of species extinctions is admittedly indirect but includes the observations that (1) empirical studies show that familial diversity correlates well with estimated species diversity over the Phanerozoic¹¹, indicating that the traditionally recognized families are capable of reflecting species-level evolutionary patterns; and (2) Monte Carlo simulations of phylogenies indicate that noncladistic taxa, with paraphyletic groups of highly variable size (including frequent monotypy), track patterns of diversity and extinction at the lineage level⁵⁻¹¹.

Families, however, are not the only sources of information on Phanerozoic extinction patterns. Although it is true that the initial evidence for periodicity rested upon stratigraphic ranges of marine families, subsequent analyses of other kinds of data have corroborated the preliminary results. Much larger data sets for marine genera^{4,9}, even after partial correction for Lagerstätten and monographic effects^{10,11}, exhibit patterns of extinction similar to the families over the Mesozoic and Cenozoic. The patterns include peaks of extinction in the upper Pliensbachian (and not in the Toarcian, containing the Posidonienschiefer), upper Tithonian (and not lower Tithonian, containing the Solnhofen), and Upper Eocene (and not Middle Eocene, containing the Monte Bolca fish beds); this indicates the data are reflecting more than mere variation in fossiliferousness of localities.

Even more importantly, most of the peaks in extinction in the familial (and the generic) data correspond to events initially recognized in detailed biostratigraphic studies, as I have repeatedly emphasized^{8,9,11,13}. Biostratigraphers are no fools, deluded by paraphyletic taxa and shifting facies. Their painstaking studies have revealed faunal discontinuities of varying magnitude in the Upper Permian, uppermost Triassic, Pliensbachian (actually lowermost Toarcian), Tithonian, Cenomanian, Maastriichtian, and Upper Eocene. These times form a periodic array with a 26 Myr spacing.

Why, then, do cladistically culled data for echinoderms and vertebrates fail to exhibit periodicity? A variety of explanations can be offered, ranging from taxonomic artefact as argued by Patterson and Smith, to an unrepresentative nature of echinoderms and vertebrates, to bias in cladistic classification. I believe the hypothesis of taxonomic artefact can be rejected quickly. Patterson and Smith's 'clade extinctions' at the top of their Fig. 1 (ref. 1) fail to show several well-documented species-level events, including the Cretaceous/Tertiary mass extinction, known since the mid-nineteenth century. This indicates something is awry with their data. It could very well be that echinoderms and vertebrates are not representative of the marine fossil record. Vertebrates have a notoriously incomplete and spotty fossil record, and the echinoderm record is hardly ideal, especially compared to the frequently fossilized molluscs. Thus, the familial record of the two groups may indeed primarily reflect fossilization noise.

But the Cretaceous/Tertiary event does seem to appear among paraphyletic taxa. (Remember that the dinosaurs, the symbol of the Cretaceous/Tertiary extinctions, constitute a paraphyletic group.) Thus, there may be a problem with cladistic classification in this context: cladistic culling may be biasing the data away from the underlying species patterns. Cladistic classification is based upon phylogenetic geometry. Thus, anything that alters the geometry may alter the classification as well. Extinction events, and especially mass extinctions, can alter phylogenetic geometries radically, and therefore cladistic classifications may become altered around these events. Polytypic, monophyletic clades thus may not approach unbiased samples of species and may no longer be useful for studying extinction time series, such as used to argue for periodicity. This suggestion should not be construed as a general criticism of cladistic taxonomy or phylogenetic systematics; cladistics is immensely valuable in many situations. It is simply an exposition of a possible limitation of the technique and a testable explanation for why Patterson and Smith failed to find periodicity, or even the well-known extinction events, in the culled sample of echinoderm and vertebrate families.

J. JOHN SEPKOSKI Jr

Department of the Geophysical Sciences,
University of Chicago,
5734 South Ellis Avenue,
Chicago, Illinois 60637, USA

1. Patterson, C. & Smith, A. B. *Nature*, **330**, 248-251 (1987).
2. Sepkoski, J. J. Jr *Contr. Biol. Geol. Milwaukee Publ. Mus.* **51**, 1-125 (1982).
3. Raup, D. M. & Sepkoski, J. J. Jr *Proc. natn. Acad. Sci. U.S.A.* **81**, 801-805 (1984).
4. Raup, D. M. & Sepkoski, J. J. Jr *Science* **231**, 833-836 (1986).
5. Sepkoski, J. J. Jr *Palaeobiology* **4**, 223-251 (1978).
6. Sepkoski, J. J. Jr *Geol. Soc. Am. Spec. Pap.* **190**, 283-289 (1982).
7. Sepkoski, J. J. Jr *Palaeobiology* **10**, 246-267 (1984).
8. Sepkoski, J. J. Jr & Raup, D. M. in *Dynamics of Extinction* (ed. Elliott, D. K.) 3-36 (Wiley, New York, 1986).
9. Sepkoski, J. J. Jr in *Patterns and Processes in the History of Life* (eds Raup, D. M. & Jablonski, D.) 277-295 (Springer, Berlin, 1986).
10. Sepkoski, J. J. Jr in *Global Bio-Events* (ed. Walliser, O.) 47-61 (Springer, Berlin, 1986).
11. Sepkoski, J. J. Jr (in preparation).
12. Sepkoski, J. J. Jr, Bambach, R. K., Raup, D. M. & Valentine, J. W. *Nature* **293**, 435-437 (1981).
13. Sepkoski, J. J. Jr & Raup, D. M. *Nature* **321**, 533 (1986).

Anaerobic production of magnetite by a dissimilatory iron-reducing microorganism

Derek R. Lovley*, John F. Stolz†, Gordon L. Nord, Jr* & Elizabeth J. P. Phillips*

*US Geological Survey, Reston, Virginia 22092, USA

†Department of Biochemistry, University of Massachusetts, Amherst, Massachusetts 01003, USA

The potential contribution of microbial metabolism to the magnetization of sediments has only recently been recognized. In the presence of oxygen, magnetotactic bacteria can form intracellular chains of magnetite while using oxygen or nitrate as the terminal electron acceptor for metabolism¹. The production of ultrafine-grained magnetite by magnetotactic bacteria in surficial aerobic sediments may contribute significantly to the natural remanent magnetism of sediments²⁻⁴. However, recent studies on iron reduction in anaerobic sediments suggested that bacteria can also generate magnetite in the absence of oxygen⁵. We report here on a sediment organism, designated GS-15, which produces copious quantities of ultrafine-grained magnetite under anaerobic conditions. GS-15 is not magnetotactic, but reduces amorphous ferric oxide to extracellular magnetite during the reduction of ferric iron as the terminal electron acceptor for organic matter oxidation. This novel metabolism may be the mechanism for the formation of ultrafine-grained magnetite in anaerobic sediments, and could

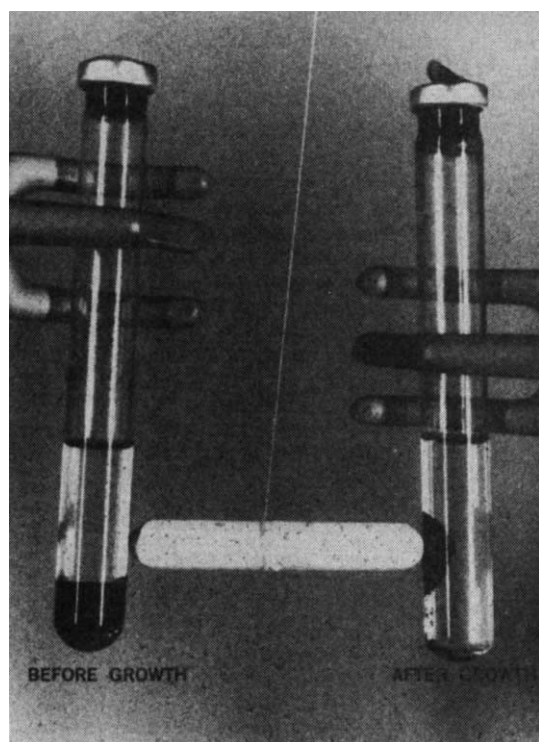


Fig. 1 Appearance of anaerobic culture medium prior to and after growth of GS-15. The brown amorphous ferric oxide at the bottom of the tube prior to growth was not attracted to the magnet. The metabolism of GS-15 resulted in the production of a large quantity of black magnetic precipitate.

account for the accumulation of magnetite in ancient iron formations and hydrocarbon deposits.

Although microbial reduction of ferric iron plays an important role in the iron geochemistry and organic matter mineralization of aquatic sediments, the pathways for organic matter metabolism and the organisms which carry out these reactions are only beginning to be identified⁵⁻⁷. Previously described bacteria which reduced ferric iron during anaerobic growth reduced iron as a minor reaction, with fermentation being the major mode of organic matter metabolism⁷. GS-15, which was isolated from sediments of the Potomac River, is the first organism known to effectively couple organic matter oxidation to ferric iron reduction during growth under anaerobic conditions. For each mole of acetate GS-15 oxidizes, two moles of carbon dioxide are produced and eight moles of ferric iron are reduced to ferrous iron. As iron serves as the terminal electron acceptor for metabolism, this reduction is termed dissimilatory iron reduction to distinguish it from the reduction of iron during microbial assimilation of intracellular iron. It was observed that dissimilatory iron reduction by GS-15 resulted in the formation of a highly magnetic black precipitate. A detailed characterization of this organism will be published separately. Here we present the analysis of the black precipitate and its implications.

The medium in which GS-15 is grown contains acetate as the sole electron donor, various major and minor minerals, and approximately 0.2 moles of ferric iron per litre in the form of amorphous ferric oxide. The amorphous ferric oxide is synthesized by neutralizing a ferric chloride solution, and has no distinct diffraction maxima detectable by powder X-ray diffraction analysis⁵. The medium contains less than 200 μ M sulphate and no reducing agents (for example, sulphide, thioglycolate, titanium). Oxygen is removed from the medium by degassing with nitrogen. Typically, the ferrous iron that is transferred with the inoculum provides an initial ferrous iron concentration in

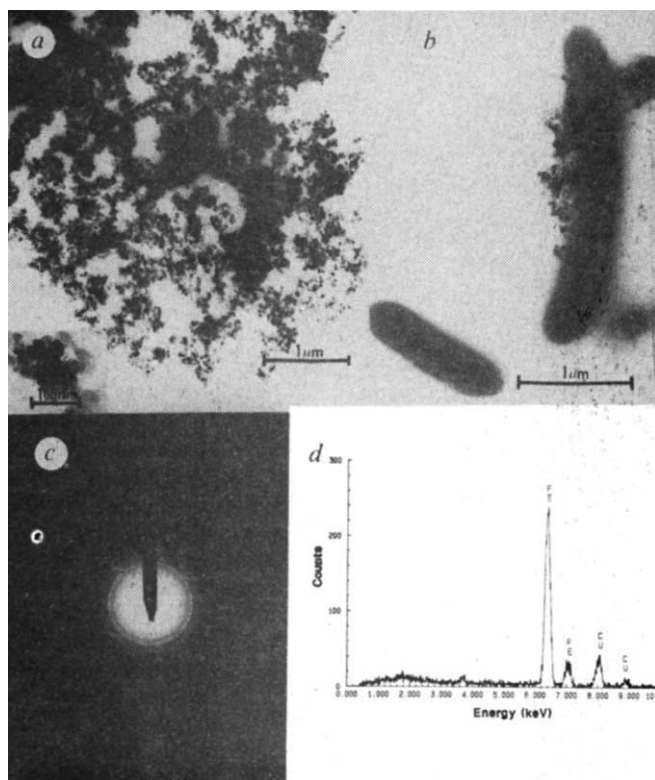


Fig. 2 Transmission electron micrographs of black precipitate and GS-15. An aggregate of the particles (*a*, bar 1 μ m) and cells of GS-15 with extracellular particles (*b*, bar 1 μ m) were observed with a JEOL 100CX operating at 80 kV. Micrographs of particles at higher magnification (inset of *a*, bar 100 nm), electron diffraction pattern (*c*) and X-ray energy dispersive analysis of the crystals (*d*) were obtained in a JEOL 200B operating at 200 kV. Copper peaks in (*d*) are from the grid.

the newly inoculated medium of 10 mM, which is more than enough to scavenge any residual oxygen.

During its growth GS-15 converted the non-magnetic brown amorphous ferric oxide to a black solid material which was strongly attracted to a magnet (Fig. 1). The black precipitate was not formed if the culture medium was not inoculated with GS-15, if the culture was incubated at temperatures too high for growth (for instance, 50 °C), or if the inoculated medium was sterilized prior to incubation. Furthermore, no black material was produced if ferrous iron (as ferrous chloride or ferrous ammonium sulphate) was added to uninoculated medium. These results indicated that the metabolism of GS-15 was necessary for formation of the black precipitate.

Examination of the black precipitate with transmission electron microscopy (TEM) revealed aggregates of small crystals ranging in size from 10 to 50 nm (Fig. 2*a*). The crystals were clearly external to the cells and were not aligned in chains (Fig. 2*b*). Further evidence for the lack of intracellular magnetite was the observation that, unlike magnetotactic bacteria, in wet mounts, GS-15 did not orient in response to an applied magnetic field. A selected area electron diffraction pattern taken from an aggregate of the crystals gave a diffuse ring pattern and a few single crystal diffraction spots (Fig. 2*c*). The lines 220, 311, 400, 422, 440, and 551 were identified by spacing and relative intensity. This diffraction pattern is characteristic of magnetite⁸. The black precipitate was not dissolved by a sodium dithionite-citrate solution which dissolves hematite, maghemite, goethite, and pyrrhotite, but not magnetite⁹. X-ray energy dispersive analysis of the crystals in the TEM detected only iron in the crystal aggregates (Fig. 2*d*).

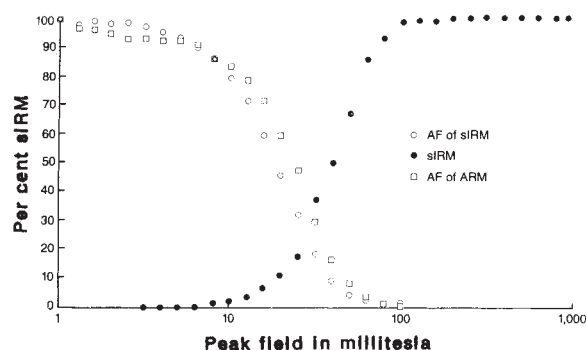


Fig. 3 Coercivity spectrum (sIRM, closed circles; AF of sIRM, open circles) and AF of ARM (squares) data for the black precipitate.

The magnetic properties of the black precipitate were determined using a SQUID magnetometer¹⁰. Anhysteretic remanent magnetization (ARM) in a biasing field of 2.5 G with subsequent alternating field demagnetization (AF) was followed by saturation isothermal remanent magnetization (sIRM) and AF demagnetization. The coercivity spectrum generated (Fig. 3) is indicative of a ferromagnetic mineral like magnetite and is similar to the coercivity spectra observed in marine sediments that contain ultrafine-grained magnetite^{2,3}. A Lowrie-Fuller test was performed by comparing the values for the median destructive field (MDF, peak field at 50% sIRM) of the AF of sIRM and AF of ARM. The MDF_{ARM} was greater than MDF_{sIRM} which suggests that the particles are magnetically single domain^{11,12}.

These results demonstrate that nonmagnetotactic bacteria can produce large quantities of ultrafine-grained magnetite under anaerobic conditions by coupling the oxidation of organic matter to the reduction of ferric iron. This has several important implications to iron geochemistry and interpretation of the sedimentary magnetic record.

Ultrafine-grained magnetite is an important carrier of remanent magnetism in sediments¹³. The recovery of magnetite crystals with morphologies characteristic of the magnetite of magnetotactic bacteria has demonstrated that magnetite production by magnetotactic bacteria contributes to the magnetization of sediments^{2-4,9,14}. However, there is substantial evidence that not all of the ultrafine-grained magnetite that is produced in sediments originates from magnetotactic bacteria. In a detailed study of the morphology of sedimentary magnetite, Chang and Kirschvink found that much of the ultrafine-grained magnetite could not be unequivocally attributed to magnetotactic bacteria and suggested that the magnetite was authigenic⁹. Karlin and co-workers have recently reported data which suggests the authigenic formation of ultrafine-grained magnetite in anaerobic sediments within the zone of iron reduction¹⁵. Although magnetotactic bacteria were considered to be the probable source, this is unlikely because they require oxygen for growth and magnetite synthesis¹. However, the data is consistent with the formation of magnetite by dissimilatory iron-reducing bacteria, such as GS-15. Further studies to ascertain the relative contributions of magnetotactic bacteria and dissimilatory iron-reducing bacteria to the magnetization of sediments are warranted.

Carbon isotope data suggests that the formation of magnetite in ancient iron formations resulted from the oxidation of organic matter coupled to the reduction of ferric oxide to magnetite^{16,17}. We propose that dissimilatory iron-reducing bacteria, similar to GS-15, were the catalysts for this reaction. Once formed, the magnetite could have persisted in the absence of significant subsequent sulphide production, because it is resistant to further microbial reduction⁵. Magnetotactic bacteria have also been proposed as a source of magnetite in these iron formations¹⁸. However, ferric iron was probably precipitated first and then

reduced under anaerobic conditions^{17,19}. This view is in accordance with metabolism of dissimilatory iron-reducing organisms but not magnetotactic bacteria. It has also been suggested that the magnetite associated with hydrocarbon deposits results from the microbial degradation of hydrocarbons²⁰. Although no bacteria have been identified yet, the finding that dissimilatory iron-reducing bacteria can produce magnetite during organic matter metabolism makes this proposed mechanism more plausible. From these considerations, it is apparent that magnetite formation by dissimilatory iron-reducing bacteria is likely to have an important influence on the geochemistry of sedimentary environments.

We thank E. Landa, D. Webster, C. Lawson, S.-R. Chang, J. Kirschvink, S. Goodwin, P. Toth, and S. Haggarty for their assistance and helpful discussions.

Received 10 July; accepted 1 October 1987.

- Blakemore, R. P., Short, K. A., Bazylinski, D. A., Rosenblatt, C. & Frankel, R. B. *Geomicrobiol. J.* **4**, 53-71 (1985).
- Stolz, J. F., Chang, S.-B. R. & Kirschvink, J. L. *Nature* **321**, 849-851 (1986).
- Stolz, J. F., Chang, S.-B. R. & Kirschvink, J. L. in *Proceedings of the 5th International Symposium on Biomineralization* (ed. Crick, R.) (Chicago University Press, Chicago, in the press).
- Chang, S.-B. R., Stolz, J. F. & Kirschvink, J. L. *Phys. Earth Planet. Int.* (in the press).
- Lovley, D. R. & Phillips, E. J. P. *Appl. Environ. Microbiol.* **52**, 751-757 (1986).
- Lovley, D. R. & Phillips, E. J. P. *Appl. Environ. Microbiol.* **51**, 683-689 (1986).
- Lovley, D. R. *Geomicrobiol. J.* **5**, 375-399 (1987).
- National Bureau of Standards, *Mongraph* **25**, Sec. 5, 31 (1967).
- Chang, S.-B. R. & Kirschvink, J. L. in *Magnetite Biomineralization and Magnetoreception in Organisms* (eds Kirschvink, J. L. et al.) 647-669 (Plenum, New York, 1985).
- Fuller, M., Goree, W. S. & Goodman, W. L. in *Magnetite Biomineralization and Magnetoreception in Organisms* (eds Kirschvink, J. L. et al.) 103-151 (Plenum, New York, 1985).
- Lowrie, W. & Fuller, M. J. *Geophys. Res.* **76**, 6339-6349 (1971).
- Johnson, H. P., Lowrie, W. & Kent, D. V. *Geophys. J. R. Astr. Soc.* **41**, 1-10 (1975).
- Kirschvink, J. L. & Chang, S.-B. R. *Geology* **12**, 559-562 (1984).
- Petersen, N., von Döbeneck, T. & Vali, H. *Nature* **320**, 611-615 (1986).
- Karlin, R., Lyle, M. & Heath, G. R. *Nature* **326**, 490-493 (1987).
- Perry, E. C., Tan, F. C. & Morey, G. B. *Econ. Geol.* **68**, 1110-1125 (1973).
- Baur, M. E., Hayes, J. M., Studley, S. A., & Walter, M. R. *Econ. Geol.* **80**, 270-282 (1985).
- Frankel, R. B. *Nature* **320**, 575 (1986).
- Walker, J. C. G. *Nature* **304**, 340-342 (1984).
- Elmore, R. D. et al. *Nature* **325**, 428-430 (1987).

Experimentally induced alteration in the polarity of developing neurons

Carlos G. Dotti & Gary A. Banker*

Department of Anatomy, Cell Biology, and Neurobiology, Albany Medical College, Albany, New York 12208, USA

Despite the great diversity of shapes exhibited by different classes of nerve cells, nearly all neurons share one feature in that they have a single axon and several dendrites. The two types of processes differ in their morphology, in their rate of growth, in the macromolecular composition of their cytoskeletons and surface membranes, and in their synaptic polarity^{1,2}. When hippocampal neurons are dissociated from the embryonic brain and cultured, they reproducibly establish this basic form with a single axon and several dendrites, despite the absence of any spatially organized environmental cues, and without the need for cell to cell contact³⁻⁷. We have cut the axons of young hippocampal neurons within a day of their development: in some cases the initial axon regenerated, but more frequently one of the other processes, which if undisturbed would have become a dendrite, instead became the axon. Frequently the stump of the original axon persisted following the transection and subsequently became a dendrite. Evidently the neuronal processes that first develop in culture have the capacity to form either axons or dendrites. The acquisition of axonal characteristics by one neuronal process apparently inhibits the others from becoming axons, so they subsequently become dendrites.

* To whom correspondence should be addressed.

Fig. 1 The effect of axonal transection on the development of polarity by hippocampal neurons in culture. Each pair of drawings illustrates the morphology of an individual hippocampal neuron before and after axotomy. The left-hand drawing in each pair shows the cell after 1 day in culture, at the time its axon was cut. The arrow indicates the site of transection. The drawing at right illustrates the same cell 20–24 h later. The cell illustrated in *a* regenerated its original axon. The cells illustrated in *b–e* exemplify cases in which the polarity of the cell was altered. In each case the new axon arose from one of the minor processes that was present when the lesion was made (asterisks). The stump of the initial axon retracted slightly from the point of lesion, but was not resorbed completely. Scale bar, 20 μ m.

Methods. Cell cultures were prepared from the hippocampi of 18 or 19 day old rat fetuses and grown on polylysine-treated glass coverslips^{4,17}. Transections were made using a micropipette mounted on a micromanipulator attached to the stage of a Zeiss IM35 inverted microscope. The transections were monitored by video recording. Typically both proximal and distal processes retracted slightly from the cut so that a clear gap of a few micrometers could be seen. Following transection, the location of lesioned cells was marked on the coverslip using a diamond object marker so that they could subsequently be relocated. The cultures were then returned to the incubator and re-examined the following day. Drawings of the cells before and after lesion were made by tracing directly from the video monitor.

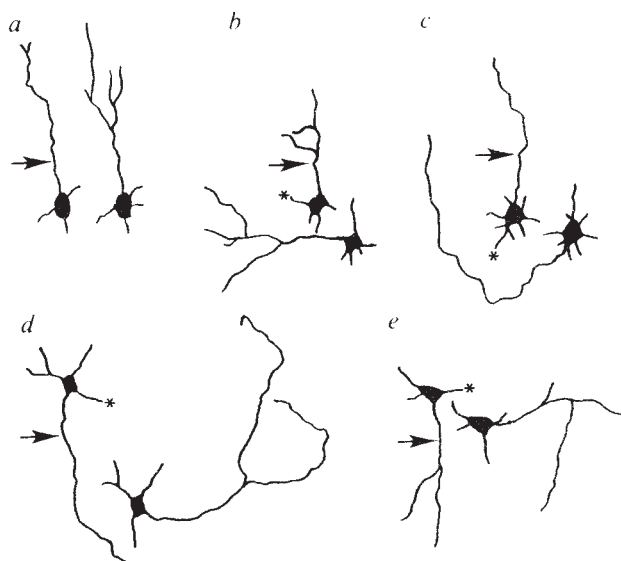
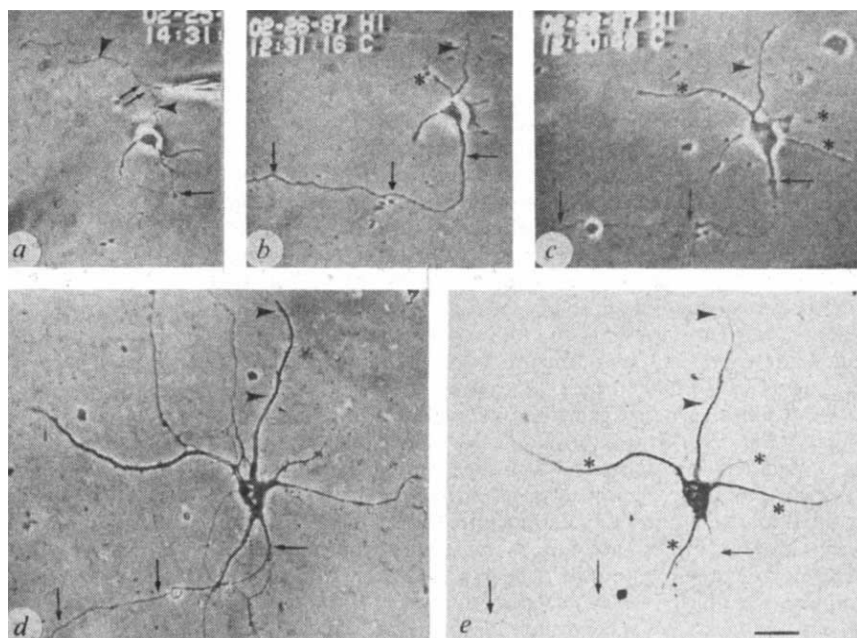


Fig. 2 Morphological and molecular characterization of the processes that develop in response to axotomy. *a*, After 1 day in culture, the axon of this cell (arrowheads) was transected with a glass micropipette (visible at right). The gap between the proximal stump and the distal axon (double arrows) is just visible in this photograph, which was taken a few seconds after lesioning. *b*, One day later a new axon had formed (arrows); it arose from one of the minor processes that was present at the time of lesion (arrow in *a*). The stump of the original axon remained (arrowhead in *b*), and a new minor process had formed near its base (asterisk in *b*). *c*, By 3 days after the lesion, the stump remaining after axotomy (arrowhead), as well as the cell's original minor processes (asterisks), had begun to elongate at the slow rate characteristic of dendrites. The newly formed axon (arrows) had elongated still further, extending far beyond the field shown. The axon from another cell entered the field from below and contacted the lesioned cell. *d, e*, Six days after the lesion the culture was fixed and immunostained to localize MAP2, using the peroxidase-antiperoxidase technique¹⁰. The cell's morphology (*d*) and the distribution of immunoreactive MAP2 (*e*) are shown in the phase contrast and bright-field photomicrographs. The dendrite that developed from the stump of the original axon (arrowheads) is quite comparable to the cell's other dendrites (asterisks) in length, shape, and content of MAP2. Note that one new dendrite formed in the interval between *c* and *d*. Most important, the new axon formed in response to axotomy is MAP2 negative, a characteristic of mature axons in culture¹⁰. The MAP2 staining of the proximal axonal segment of the lesioned cell is typical of normal cells as well. Scale bar, 20 μ m.



Previous work has shown that hippocampal neurons in culture acquire their characteristic form by a stereotyped sequence of developmental events³. Initially several short, minor processes arise from the cell. Then 4–12 h later, one of these begins to elongate rapidly. It becomes the cell's axon. After a few days, the remaining minor processes begin to elongate to become the cell's dendrites. Two observations suggest that processes may not be specified as axonal or dendritic at the time they first appear, and that any of the minor processes could become either axons or dendrites³. First, the minor processes initially are indistinguishable in the light microscope and in their behaviour in time-lapse video recordings. Second, minor processes occasionally compete with one another to become the cell's axon. First one grows at the rapid rate characteristic of axons,

then the other. Ultimately one becomes the definitive axon and the other retracts to become indistinguishable from the other minor processes.

To test the possibility that the axonal or dendritic fate of minor processes is not rigidly specified, we transected the newly formed axons of hippocampal neurons in culture. One-day-old cells with a single, clearly defined axon were selected and their axons cut with a micropipette^{8,9}. Responses from the cells are illustrated in Fig. 1, and the results summarized in Table 1. Within a day, nearly all the cells that survived the lesions reacquired the morphology typical of polarized hippocampal neurons, having a single axon and several minor processes. Cells without axons were rare (6%). In some cases the original axon regenerated (Fig. 1*a*). More frequently, however, one of the

Table 1 The response of hippocampal neurons to transection of their axons after 1 day in culture

	Number of cases
Cells lesioned	103
Cells surviving 24 h	81
Cells redeveloping axons within 24 h	76
Original axon regenerated	26 (34%)
New axon arose from a minor process	42 (55%)
New axon arose de novo	3 (4%)
Two new axons arose	5 (7%)

undamaged minor processes took on axonal characteristics, while the stump of the cut axon remained short, like the minor processes of the cell (Fig. 1b-f). Occasionally a new axon formed *de novo* rather than from a minor process, and sometimes cells developed two axons, but a few normal hippocampal cells also develop two axons at this stage (K. A. Gostin and G.A.B., unpublished observations).

The normal events in the development of hippocampal neurons that occur after axon initiation are stereotyped: the axon itself continues to elongate and frequently branches extensively, then after a delay of several days, the minor processes elongate at a much slower rate to become dendrites. At the end of a week in culture, the axonal and dendritic domains have acquired their characteristic morphology and have a distinct set of molecular constituents^{3,4,6}. In particular, the microtubule-associated protein MAP2 is restricted to the cell body and dendrites¹⁰.

These observations of normal hippocampal neurons raise additional questions concerning the alterations in polarity induced by axotomy. Do the new axons persist and differentiate normally? Do regionally localized constituents like MAP2 become appropriately compartmentalized in lesioned cells? What is the fate of the stump of the initial axon? We followed cells whose axons had been cut after 1 day in culture for the next 6 days. An example is shown in Fig. 2: one day after the original axon was cut, one of the minor processes had become the new axon (Fig. 2b), the stump of the original axon persisted and a new branch had formed near its base. After three days the initial minor processes and the axon stump had begun to elongate (Fig. 2c). The new axon had meanwhile extended further and established several branches (not included in the field shown in Fig. 2c). Six days after the lesion, the culture was fixed and immunostained to localize MAP2 (Fig. 2d and e). At this stage, the axon stump resembled the other dendrites in length and tapering profile. The new axon that formed in response to axotomy was MAP2 negative, as are normal hippocampal axons at this stage of development. The dendrite that had formed from the original axon showed MAP2 immunoreactivity, like the cell's other dendrites and like the dendrites of normal hippocampal neurons. In all, we followed 12 cells for 6 days after axotomy, then immunostained them to localize MAP2. Comparable results were observed in all cases.

Transecting the axons of mature neurons can alter some aspects of their normal polarity, especially if the lesion occurs near the cell body. In some cases new axons sprout from the dendrites, the dendritic membrane can acquire physiological properties normally characteristic of axons, the dendrites may even acquire a myelin sheath¹¹⁻¹⁶. Our results show that transecting the axons of developing neurons in culture alters their subsequent development in a much more fundamental way. A process that would normally become a dendrite can become the axon, and the axon can become a dendrite. It is as if the lesion returns the cell to an earlier stage of development, when each of its processes has a chance of becoming the axon. If this interpretation is correct, it remains to be determined how the assumption of axonal characteristics by one process prevents the others from becoming axons.

Remarkably, despite this profound change in the specification of individual processes, the fundamental polarization of the neuron develops quite normally: the cell still establishes a single axon and several dendrites. Study of the cellular reorganization that accompanies the lesion-induced alteration of polarity, in combination with studies of the normal process of polarization, may provide a valuable new paradigm for determining how neuronal polarity is established.

We thank Dr L. I. Binder for the gift of MAP2 antibody and Colleen Fitzgerald for technical assistance. This work was supported by the NIH; C. Dotti was the recipient of a Fogarty International Fellowship from the NIH.

Received 6 August; accepted 2 October 1987.

1. Peters, A., Palay, S. & Webster, H. de F. *The Fine Structure of the Nervous System: The Neurons and Supporting Cells* (Saunders, Philadelphia, 1976).
2. Banker, G. A. & Waxman, A. B. in *Intrinsic Determinants of Neuronal Form and Function* (eds Lasek, R. J. & Black, M. B.) (Liss, New York, 1987).
3. Dotti, C. G., Sullivan, C. A. & Banker, G. A. *J. Neurosci.* (in the press).
4. Bartlett, W. P. & Banker, G. A. *J. Neurosci.* 4, 1944-1953 (1984).
5. Bartlett, W. P. & Banker, G. A. *J. Neurosci.* 4, 1954-1965 (1984).
6. Caceres, A., Banker, G. A., Steward, O., Binder, L. & Payne, M. R. *Devl Brain Res.* 13, 314-318 (1984).
7. Banker, G. A. & Mazurkiewicz, J. E. *Soc. Neurosci. Abstr.* 8, 696 (1982).
8. Shaw, G. & Bray, D. *Expl Cell Res.* 104, 55-63 (1977).
9. Wessells, N. K., Johnson, S. R. & Nuttall, R. P. *Expl Cell Res.* 117, 335-345 (1978).
10. Caceres, A., Banker, G. A. & Binder, L. *J. Neurosci.* 6, 714-722 (1986).
11. Hall, G. F. & Cohen, M. J. *Science* 222, 518-521 (1983).
12. Linda, H., Risling, M. & Cullheim, S. *Brain Res.* 358, 329-333 (1985).
13. Eccles, J. C., Libet, B. & Young, B. R. *J. Physiol., Lond.* 143, 11-40 (1958).
14. Kuno, M. & Llinas, R. *J. Physiol., Lond.* 210, 807-821 (1970).
15. Titmus, M. J., Faber, D. S. & Zottoli, S. J. *J. Neurophysiol.* 55, 1424-1439 (1986).
16. Hamori, J., Lakos, I. & Mezey, E. *J. Hirnforsch.* 21, 391-407 (1980).
17. Banker, G. A. & Cowan, W. M. *Brain Res.* 126, 397-425 (1977).

Interaction between CD4 and class II MHC molecules mediates cell adhesion

Carolyn Doyle & Jack L. Strominger

Department of Biochemistry and Molecular Biology,
Harvard University, 7 Divinity Avenue, Cambridge,
Massachusetts 02138, USA

The CD4 glycoprotein is expressed on T-helper and cytotoxic lymphocytes which are restricted to class II major histocompatibility complex (MHC) antigens on target cells¹⁻⁵. Antibody inhibition studies imply that CD4 acts to increase the avidity of effector-target cell interactions⁶⁻⁸. These observations have led to the speculation that CD4 binds to a monomorphic class II antigen determinant, thereby augmenting low affinity T-cell receptor-antigen interactions⁹⁻¹¹. However, no direct evidence has been presented indicating that CD4 and class II molecules interact. To address this issue, we have used a vector derived from simian virus 40 (SV40)¹² to express a complementary DNA (cDNA) encoding the human CD4 glycoprotein¹³. When CV1 cells expressing large amounts of the CD4 protein at the cell surface are incubated with human B cells bearing MHC-encoded class II molecules, they are bound tightly to the infected monolayer, whereas mutant B cells which lack class II molecules fail to bind. Furthermore, the binding reaction is specifically inhibited by anti-class II and anti-CD4 antibodies. Thus, the CD4 protein, even in the absence of T-cell receptor-antigen interactions, can interact directly with class II antigens to function as a cell surface adhesion molecule.

The SV40-derived recombinant vector used to express the CD4 protein is essentially identical to that used previously to produce large amounts of the influenza haemagglutinin glycoprotein in CV1 cells¹². The cDNA sequences are transcribed under the control of the SV40 late promoter and replace the viral sequences which normally encode the capsid proteins required to package the lytic virus. The CD4 cDNA sequences are inserted between the *Hpa*II (346-base pair (bp) and the

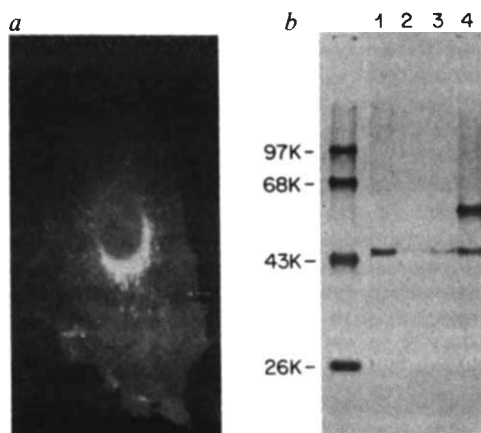


Fig. 1 Indirect immunofluorescent labelling of CV1 cells infected with the SV40-CD4 recombinant virus. *a*, CV1 cells were typically prepared for infection by plating in 60mm tissue culture dishes at a density of approximately $4-5 \times 10^5$ cells in Dulbecco's modified Eagle's medium (DMEM) and 10% foetal calf serum (FCS). The medium was removed 8–10 h later and cells were incubated with 0.25 ml of an SV40-CD4 (or an SV40-HA) high-titre virus stock (multiplicity of infection of 10 plaque-forming units per cell) for 1.5–2 h at 37 °C. 5 ml of DMEM-10% FCS was then added and virus was grown at 37 °C for 36–60 h at which time the experiments were performed. For immunofluorescent labelling, SV40-CD4 infected CV1 cells grown on coverslips were fixed using 3.7% formaldehyde in phosphate-buffered serum and permeabilized with 0.1% Triton X-100 as previously described¹⁷. Indirect immunofluorescent staining was carried out using a CD4-specific monoclonal antibody (anti-Leu3A, Becton-Dickinson) followed by goat anti-mouse IgG conjugated with rhodamine (Cappel Laboratories). *b*, SDS-polyacrylamide gel electrophoresis of CD4 protein immunoprecipitated from extracts of CV1 cells infected with SV40-CD4 viral vectors. SV40-infected cells were immunoprecipitated with anti-Leu3A (lane 1). SV40-CD4 infected cells were immunoprecipitated with OKT8 (lane 2), OKT4 (lane 3), and anti-Leu3A (lane 4).

Methods. CV1 cells, infected for 42 h with wild-type SV40 or recombinant SV40-CD4 virus stocks, were incubated for 30 mins in DMEM lacking methionine and labelled for one hour with [³⁵S]methionine (100 μ Ci per 60 mm plate; 1,000–3,000 Ci mmol⁻¹, Amersham) in 0.25 ml DMEM lacking methionine. Cells were lysed in 1% Nonidet P-40, 10 mM Tris (pH 8.0) and aliquots of the cell extracts were immunoprecipitated in NET-GEL (150 mM NaCl, 5 mM EDTA, 10 mM Tris (pH 8.0), 0.25% gelatin) using a nonspecific (OKT8) or CD4-specific monoclonal antibodies anti-Leu3A (4 μ g) and OKT4 (5 μ g) followed by incubation with rabbit anti-mouse serum and Protein-A Sepharose. The precipitated polypeptides were then washed extensively in NET-GEL+0.5M NaCl, NET-GEL+1% NP40 and 0.1% SDS and finally in 10 mM Tris-HCl containing 0.1% NP40. The proteins were then solubilized by boiling in Laemmli sample buffer²⁵ and separated on discontinuous SDS-polyacrylamide gels (0.1% SDS; 10% polyacrylamide; 30:0.8/acrylamide:Bis) and visualized by fluorography at –70 °C using Kodak XAR-5 film.

*Bam*HI (2,533-bp) sites with conversion of the *Hpa*II site to *Sal*I. The recombinant viral genome, along with a defective helper virus¹⁴, is introduced into the recipient fibroblasts by transfection using DEAE-Dextran¹⁵ and chloroquine as facilitators¹⁶ as previously described¹⁷. Serial passage on freshly growing monolayers provides a high titre virus stock which can subsequently be used to produce large amounts of protein in more than 90 % of the cells.

To establish that the protein expressed from the recombinant vector was authentic in its structure, antigenicity and cellular localization, immunoprecipitation and staining by immunofluorescence using CD4-specific monoclonal antibodies, anti-Leu3A and OKT4, were performed. The immunofluorescent labelling pattern obtained is shown in Fig. 1*a*. The strong perinuclear staining is typical of that seen when glycoproteins synthe-

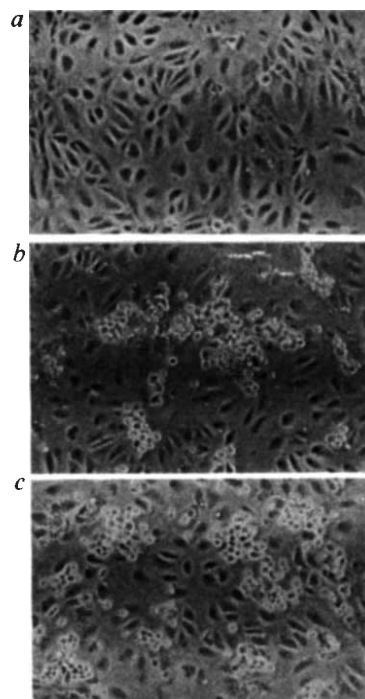


Fig. 2 Binding of Raji B lymphocytes to monolayers of CV1 cells infected with SV40 viruses. *a*, Wild-type SV40-infected cells. *b*, SV40-HA infected cells. *c*, SV40-CD4 infected cells.

Methods. At 42 h, the medium was removed from plates of CV1 cells infected with either wild-type or recombinant viruses. The monolayers were then incubated for one h at 37 °C with $1-2 \times 10^7$ B cells in 0.4 ml DMEM-10% FCS. The plates were then washed extensively with DMEM to remove non-adherent lymphocytes prior to inspection. Bound lymphocytes remained attached for more than 6–12 hours after washing at which time the monolayer began to detach due to cytopathic effects of the lytic virus cycle.

sized in the endoplasmic reticulum are transported to the cell surface via the Golgi apparatus. A diffuse surface staining (Fig. 1*a*) is apparent on approximately 25 % of the infected cells and probably corresponds to those cells in the population which were infected early during the course of the experiment and which are also likely to be expressing the most protein at the cell surface. This population expresses approximately $10-15 \times$ the number of CD4 molecules at the cell surface, as does the human T lymphocyte cell line, HPB-ALL, as determined by cytofluorographic analyses using the anti-Leu3A monoclonal antibody (data not shown). Preliminary evidence suggests that the yield of CD4 purified from these cells is in the range of 1–10 μ g per dish ($5-6 \times 10^6$ cells per 10cm plate). Extracts of cells infected with the recombinant virus and labelled with [³⁵S]methionine contained a protein with a relative molecular mass of 55K, that was specifically immunoprecipitated with both the OKT4 and the anti-Leu3A mAbs (Fig. 1*b*, lanes 3 and 4).

A lymphocyte binding assay was used to study CD4-class II interactions (Figs 2 and 3). CV1 cells were infected with wild-type SV40 virus or with a recombinant virus producing either the influenza HA or the CD4 glycoprotein. At 42 hours post-infection, the infected monolayers were incubated with B cells and examined. The results of a typical experiment using the B lymphoblastoid cell line, Raji, which expresses high levels of class II antigen(s) is shown in Fig. 2 and quantified in Fig. 3*a*. B lymphocytes, like erythrocytes, adhere to CV1 cells expressing the HA via sialic acid receptor sites present on that molecule (Fig. 2*b*). A similar binding phenomenon was observed when Raji cells were incubated with CV1 cells infected with the SV40-CD4 recombinant (Fig. 2*c*) but not with cells infected with a wild type SV40 virus (Fig. 2*a*). Similar experiments were

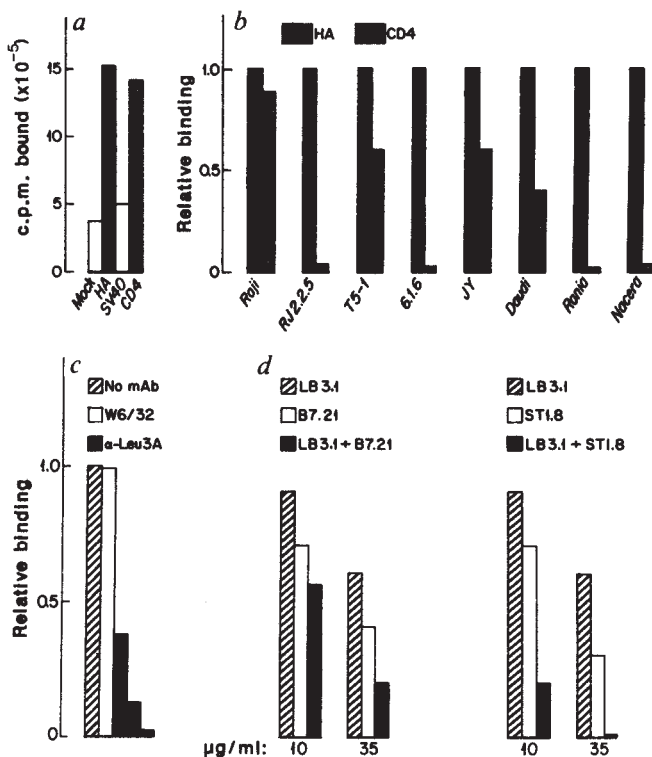


Fig. 3 Binding of lymphocytes to CV1 cells infected with SV40, SV40-HA and SV40-CD4 virus stocks. **a**, Binding of radiolabelled Raji lymphocytes to mock, SV40, SV40-HA and SV40-CD4 infected monolayers. **b**, Binding of class II-positive (Raji, T5-1, JY, Daudi) and class II-negative (RJ2.2.5, 6.1.6, Rania and Nacera) B lymphocytes to monolayers of CV1 cells expressing HA (solid bars) or CD4 (hatched bars). **c** and **d**, Antibody inhibitions of lymphocyte binding to SV40-CD4 infected cells. Monoclonal antibodies used were, for panel **c**: W6/32 (anti-class I), 100 µg ml⁻¹; anti-Leu3A (CD4-specific), 10, 50, 80 µg ml⁻¹. For panel **d**: LB3.1 (DR-specific), 10 and 35 µg ml⁻¹; and B7.21 and ST1.8 (DP-specific), 10 and 35 µg ml⁻¹.

Methods. **a**, Raji B lymphocytes ($1-2 \times 10^7$ cells) were incubated in methionine-free DMEM containing 200 µCi ml⁻¹ [³⁵S]methionine for one hour. Binding of the radiolabelled Raji cells to infected monolayers was carried out as described above (see legend Fig. 2), followed by extensive washing. Monolayers and bound lymphocytes were then removed from dishes by trypsinization, dissolved in scintillation fluid and radioactivity (c.p.m. bound) was determined. **b**, The binding of radiolabelled B lymphocytes to monolayers of CV1 cells infected with either the SV40-haemagglutinin or the SV40-CD4 recombinant viruses was quantified as above. Results represent the average of several determinations. However, in this case, the background (c.p.m. bound to wild-type SV40-infected monolayers) was subtracted and binding to SV40-HA-infected cells was normalized for comparisons among cell lines. **c**, CV1 cells infected with the SV40-CD4 recombinant virus and radiolabelled Raji B cells were preincubated with the W6/32 (100 µg ml⁻¹) or anti-Leu3A monoclonal antibodies (10, 50, 80 µg ml⁻¹) for two hours after which time the binding assay was performed. **d**, Fab' fragments of the monoclonal antibodies were employed. They were prepared by pepsin treatment (Sigma, 10 mg ml⁻¹ in 0.1 M sodium acetate, pH 4.5) of IgG for 16 h at 37 °C followed by the addition of 2M Tris base. The pFc' portion of the molecule(s) was removed by incubation with 0.2 ml of 10% protein A-Sepharose. Inhibition of binding was carried out as above.

trast, immunoselected mutants of Raji and T5-1, which do not express class II antigens (RJ2.2.5, ref. 18, and 6.1.6, ref. 19, respectively) failed to bind to CD4-producing cells. Furthermore, other class II negative B cell lines (Rania and Nacera), derived from patients afflicted with severe combined immunodeficiency disease (SCID)²⁰, behaved similarly (Fig. 3b). Importantly, these mutant cell lines, although similar in phenotype, were isolated independently. RJ2.2.5 and 6.1.6. were generated by mutagenesis *in vitro* whereas the cell lines isolated from SCID patients arose spontaneously. Therefore, these cell lines are unlikely to contain a common pleiotropic mutation affecting loci other than class II. The Daudi cell line²¹, which expresses class II but not class I antigens (because of a defect in $\beta 2$ -microglobulin expression), bound to cells expressing the CD4 molecule. Thus, the expression of class II antigens on the surface of B lymphocytes correlates with the ability of those cells to attach to monolayers of cells which express the CD4 glycoprotein. Notably, when binding was quantified using radio-labelled lymphocytes (Fig. 3b), higher levels of B lymphocyte binding to cells expressing the CD4 molecule was coincident with higher levels of class II expression (Raji > T5-1 > JY > Daudi), as determined by FACS analysis (data not shown).

The results of antibody inhibition studies confirmed the specificity of the CD4-class II interaction. The binding of class II positive B lymphocytes was completely inhibited by saturation with the anti-Leu3A monoclonal (Fig. 3c) or by a mixture of class II-specific antibodies (Fig. 3d). When either the DR-specific monoclonal antibody, LB3.1 (refs 22, 23) or DP-specific antibodies, B7.21 (refs 22, 24) or ST1.8 (J. C. Gorga, personal communication), were used alone, binding was only partially reduced (40–60%). Similar partial inhibitions were achieved using lower concentrations of a mixture of these class II monoclonal antibodies (see Fig. 3d). In contrast, a class I-specific antibody, W6/32 (Fig. 3c), failed to inhibit the binding activity and the binding of Raji cells to CV1 cells expressing the HA could not be blocked with any of the reagents used in these studies.

Taken together, these data demonstrate that the adhesion of lymphocytes bearing class II antigens to fibroblasts expressing the CD4 glycoprotein is a consequence of direct interactions between these two molecules. Furthermore, because adhesion mediated by these molecules is observed in a heterologous system (in the absence of any T-cell receptor-antigen interactions), the binding of CD4 and class II molecules can be independent of such interactions. We presume that the high level of CD4 expression obtained in this study was crucial for demonstrating this high affinity, stable interaction. At more physiological levels of expression, it is likely that CD4 and class II antigens help to mediate low-affinity, transient interactions among lymphocytes. Under such circumstances, these molecules must work in parallel with other specific and accessory adhesion molecules to allow functional cell-cell interactions.

We are grateful to Drs R. S. Accolla, D. Pious, C. Griscelli and B. Lisowska-Grospierre for providing mutant and patient cell lines, to Dr R. Axel for the CD4 cDNA clone and to Drs M.-J. Gething and J. Sambrook for the SV40 vector system. We also thank Drs M. S. Krangel, J. C. Gorga and D. R. Johnson for advice and critical reading of the manuscript and S. Liang for assistance in initial experiments. This work was supported by NIH and the American Foundation for AIDS Research to J.L.S. and a National Research Service Award to C.D.

Received 24 September; accepted 5 October 1987.

then performed with a number of different B lymphoblastoid cell lines, some of which express class II antigens, some of which do not. All B cells examined displayed binding to monolayers expressing the HA. However, only those B cells which express class II antigens, that is, Raji, Daudi, JY, and T5-1, bound to cells producing the CD4 glycoprotein. In con-

- Reinherz, E. L. & Schlossman, S. F. *Cell* **19**, 821–827 (1980).
- Engleman, E. G., Benike, C. J., Grumet, F. C. & Evans, R. L. *J. Immunol.* **127**, 2124–2129 (1981).
- Biddison, W., Rao, P., Talle, M. A., Goldstein, G. & Shaw, S. *J. exp. Med.* **156**, 1065–1076 (1982).
- Krensky, A. M., Reiss, C. S., Mier, J. W., Strominger, J. L. & Burakoff, S. J. *Proc. natn. Acad. Sci. USA* **79**, 2365–2369 (1982).
- Meuer, S. C., Schlossman, S. F. & Reinherz, E. *Proc. natn. Acad. Sci. USA* **79**, 4395–4399 (1982).

6. Wilde, D. B., Marrack, P., Kappler, J., Dyalnas, D. P. & Fitch, F. W., *J. Immun.* **131**, 2178-2183 (1983).
7. Rogozinski, L., et al. *J. Immun.* **132**, 735-739 (1984).
8. Swain, S. L., Dyalnas, D., Fitch, F. W., & English, M. *J. Immun.* **132**, 118-123 (1984).
9. Marrack, P., et al. *J. exp. Med.* **158**, 1077-1092 (1983).
10. Greenstein, J. L., Kappler, J., Marrack, P. and Burakoff, S. J. *J. exp. Med.* **159**, 1213-1224 (1984).
11. Greenstein, J. L., Malissen, B. & Burakoff, S. J. *J. exp. Med.* **162**, 369-374 (1985).
12. Gething, M. J. & Sambrook, J. *Nature* **293**, 620-625 (1981).
13. Maddon, P. J., et al. *Cell* **42**, 93-104 (1985).
14. Pipas, J. M., Adler, S. P., Peden, K. W. C. & Nathans, D. *Cold Spring Harb. Symp. quant. Biol.* **44**, 285-291 (1980).
15. McCutchan, J. H. & Pagano, J. S. *J. natn. Cancer Inst.* **41**, 351-357 (1969).
16. Luthman, H. & Magnusson, G. *Nucleic Acids Res.* **11**, 1295-1308 (1983).
17. Doyle, C., Roth, M. G., Sambrook, J. & Gething, M. J. *J. Cell Biol.* **100**, 704-714 (1985).
18. Accolla, R. S. *J. exp. Med.* **157**, 1053-1058 (1983).
19. Gladstone, P. & Pious, D. *Nature* **271**, 459-461.
20. Lisowska-Grospierre, B., et al. *J. clin. Invest.* **76**, 381-385 (1985).
21. De Preval, C., & Mach, B. *Immunogenetics* **17**, 133-140 (1983).
22. Gorga, J. C., Horejsi, V., Johnson, D. R., Raghupathy, R. & Strominger, J. L. *J. biol. Chem.* (in the press).
23. Gorga, J. C., Knudson, P. J., Foran, J. A., Strominger, J. L. & Burakoff, S. J. *cell. Immun.* **103**, 160-173 (1986).
24. Watson, A. J., DeMars, R., Trowbridge, I. S. & Bach, F. H. *Nature* **304**, 358-361 (1983).
25. Laemmli, U. K. *Nature* **227**, 680-685 (1970).

Prevention of vaccinia virus infection in immunodeficient mice by vector-directed IL-2 expression

Charles Flexner, Ambros Hügin & Bernard Moss

Laboratory of Viral Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA

Recombinant vaccinia viruses have been proposed as live vaccines against a variety of infectious diseases, including AIDS (acquired immune deficiency syndrome)¹. Objections have been concerned primarily with side effects of the vaccinia virus vector itself. Recently it has been shown that inactivation of the vaccinia virus thymidine kinase gene or deletion of certain other non-essential genes is associated with a marked reduction in pathogenicity². Nevertheless, the ability of vaccinia virus to produce a progressive infection in immunodeficient individuals remains a most serious problem. Indeed, an incident of this type in a vaccinated man seropositive for human immunodeficiency virus was recently reported³. We have used immunodeficient athymic nude mice to establish a model of disseminated vaccinia virus infection, and to demonstrate a novel approach to virus attenuation which involves insertion of a gene encoding human interleukin-2 into the genome of vaccinia virus vectors.

Human interleukin-2 (IL-2) is secreted by helper T (CD4) lymphocytes in response to antigenic stimulation and is important for the activation of specific and non-specific immune responses. The human gene encoding this small glycoprotein has been identified, sequenced, and expressed in bacterial and eukaryotic expression vectors⁴⁻⁶, including vaccinia virus⁷. Recently, IL-2 has been used in experimental animals and in man to stimulate immunity to infectious pathogens and malignant tumours⁸⁻¹⁰. In addition, local administration of IL-2 with antigen can enhance humoral immunity¹¹. The object of this study was to determine whether a recombinant vaccinia virus that expresses human IL-2 would stimulate an enhanced immune response and exhibit a more attenuated phenotype, particularly in immunodeficient animals.

Initial experiments were devoted to developing a mouse model for disseminated vaccinia virus infection. Athymic nude (nu/nu) mice, which have severely depressed numbers of T lymphocytes, were inoculated intraperitoneally with various doses of vaccinia virus. The WR strain of vaccinia virus, which has been passaged repeatedly in mouse brain, was chosen because of its virulence for mice. Death occurred in a dose-dependent manner (Fig. 1) either from wild-type virus (data not shown) or from a recom-

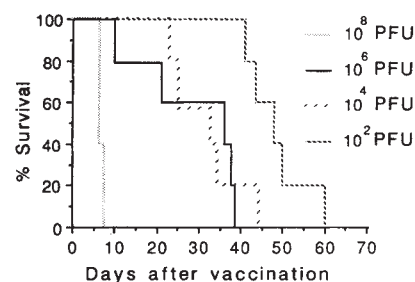


Fig. 1 Dose-dependent mortality after injection with vaccinia virus vector vTFCLZ-1 in athymic nude mice. Six-week-old male athymic nude (nu/nu) mice were injected i.p. with 0.2 ml vaccinia virus diluted in PBS at day 0, five mice for each dose of virus. Animals in all experiments were housed in adjacent cages in the same room, in laminar flow hoods under sterile conditions. Construction of vector vTFCLZ-1 is described in the text; virus stocks were purified, characterized and titred on BCS-1 cells as described previously^{14,20}.

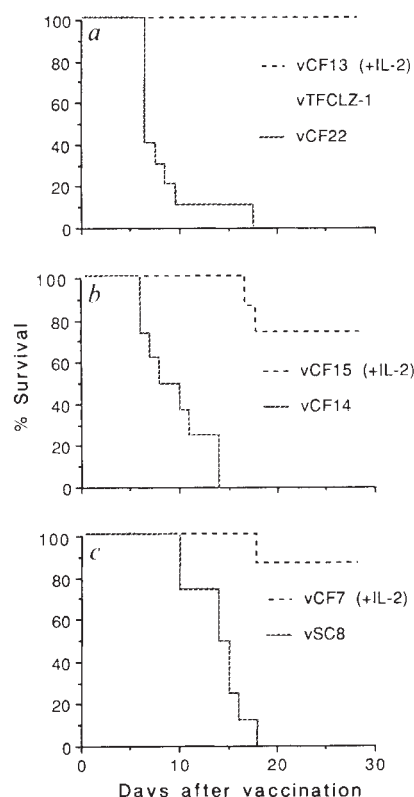
binant virus vTFCLZ-1 obtained from T. Fuerst). The recombinant virus contained the complete *Escherichia coli* β -galactosidase (*lacZ*) gene under the control of the vaccinia P11 promoter and inserted into a 400 base pair (bp) *Bam*HI deletion in the *Hind*III C region of the vaccinia genome¹². Mice that received 10^8 plaque-forming-units (PFU) of virus developed pox lesions in the skin within 3 days, exhibited weakness and cachexia, and died within 8 days. The onset of illness was delayed in mice that received lower doses of virus, and those that received only 100 PFU survived for as long as 60 days and had no symptoms until one to two weeks before death (Fig. 1).

We constructed recombinant vaccinia viruses that contain the human IL-2 gene⁴ (vCF13) or the influenza A/PR/8/34 haemagglutinin gene (vCF22) under control of the vaccinia virus P7.5 promoter and inserted into the same locus as the *lacZ* gene in vTFCLZ-1 (see Fig. 2 legend). CV-1 cells infected with vCF13 secreted a glycoprotein of relative molecular mass (M_r) 15,000 (15K) that was immunoprecipitated with rabbit antiserum raised against human IL-2. The media covering 2×10^6 CV-1 cells infected with 10 PFU per cell of vCF13 contained 32.7 units per ml IL-2 bioactivity after a 48 h infection, as measured by proliferation of murine IL-2-dependent T lymphocytes (CTL 53 (ref. 5)) compared to a human Jurkat cell line IL-2 standard. IL-2 bioactivity was detectable in the serum of BALB/c mice six days after intraperitoneal (i.p.) injection of 10^7 PFU of vCF13, but not in mice given the control vector vTFCLZ-1. Nu/nu mice injected with the control viruses vTFCLZ-1 or vCF22 developed pox lesions in the skin within 3 days and all but one were dead in 9 days (Fig. 2a). In contrast, all of the nu/nu mice injected with vCF13 remained healthy for at least 60 days (Fig. 2a). BALB/c nude mice given 10^8 PFU of vTFCLZ-1 or vCF13 showed the same pattern of mortality as athymic nude mice, except that death with vTFCLZ-1 occurred 10-14 days after inoculation, and skin lesions, though present, were less prominent.

Titres of virus in the spleens and livers of nu/nu mice inoculated with 10^7 PFU control virus vTFCLZ-1 or IL-2 virus vCF13 were measured in groups of three mice killed at 3, 7, 10 and 15 days after inoculation. Mean titres of vTFCLZ-1 were quite high, about 2×10^6 PFU per 100 mg tissue, until death by day 12, whereas the titres in mice inoculated with 10^7 PFU of IL-2-expressing virus vCF13 were never higher than 2×10^2 PFU per 100 mg tissue; virus was undetectable after day 3 in the livers of vCF13-injected mice, and was undetectable in spleens after day 10. Investigation into the mechanism of elimination of IL-2-expressing vaccinia virus revealed similar anti-vaccinia IgM titres by enzyme-linked immunosorbent assay (ELISA) and natural killer cell activity (determined by specific lysis of ⁵¹Cr-labelled YAC-1 cells 3, 7, and 10 days after injection of 10^7 PFU

Fig. 2 Mortality in nu/nu mice injected with recombinant vaccinia virus vectors expressing either *E. coli* β -galactosidase (β -gal), or β -gal plus human interleukin-2. Vectors are grouped by site of insertion of β -gal and *IL-2* genes in the vaccinia genome: *a*, vectors with inserts in the vaccinia *HindIII* *C* region near the left hand terminus of the genome; *b*, inserts in the vaccinia haemagglutinin, about 130 kilobase pairs (kb) to the right of *HindIII* *C*; *c*, inserts in the vaccinia thymidine kinase (*TK*), about 60 kb to the right of *HindIII* *C*. Results shown are for groups of 10 (*a*) or 8 (*b*, *c*) six-week-old male athymic nude mice injected i.p. with 10^8 PFU of purified recombinant vaccinia vector in 0.2 ml PBS.

Methods. Recombinant plasmid pCF11 was constructed to direct insertion of foreign genes for expression in the *Bam*HI deletion of the *HindIII* *C* section of the vaccinia genome contained in plasmid pMH5/1 (ref. 12); pCF11 contains the complete *E. coli lacZ* gene under the control of the vaccinia P11 promoter, and the P7.5 promoter with a unique *Sma*I site for insertion of foreign genes (C.F., manuscript in preparation). Recombinant vaccinia viruses that contain either the human *IL-2* gene⁴ (vCF13) or the influenza haemagglutinin gene (vCF22) under control of the vaccinia virus P7.5 promoter, inserted into the same locus as the *lacZ* gene in vTFCLZ-1, were constructed by transfection of CV-1 cells infected with the WR strain of vaccinia virus with a plasmid pCF13, containing the *IL-2* gene inserted into the *Sma*I site of pCF11, or plasmid pYF2, in which the influenza A/PR/8/34 haemagglutinin gene was inserted into the *Sma*I site of pCF11 (C.F., manuscript in preparation). Recombinants vCF14 and vCF15 have the *lacZ* gene alone or with the human *IL-2* gene respectively, inserted into the *Sma*I site of pVY6, a plasmid identical to pCF11, but directing insertion into the centre of the vaccinia haemagglutinin gene¹³ (C.F. and V. Yuferov, unpublished data). vSC8 (ref. 14) and vCF7 (ref. 7) have the *lacZ* gene alone or with the *IL-2* gene respectively, inserted into the middle of the vaccinia *TK* gene. vCF7 and vCF15 were shown to express *IL-2* at levels equal to or higher than that of vCF13 (see text). Insertional inactivation of the *TK* gene was determined by plaque formation in TK⁻ 143 cells in the presence of 25 μ g ml⁻¹ of 5-bromodeoxyuridine, and inactivation of the haemagglutinin was demonstrated by the inability of virus plaques on BSC-1 cells to agglutinate washed chicken erythrocytes²¹. Recombinant vectors were characterized, purified, and titred on BSC-1 cells as described^{14,20}.



virus) in nu/nu mice infected with vTFCLZ-1 (*IL-2*⁻) or vCF13(*IL-2*⁺). Anti-vaccinia IgG titres (by ELISA) and cytotoxic T lymphocyte activities (determined in BALB/c nude mice) were low or undetectable in both cases, as expected for nu/nu mice (data not shown).

Other recombinant vaccinia viruses were made to see if deletion of certain vaccinia virus genes had an attenuating effect in the nu/nu mouse model, and to establish that the site of insertion of the *IL-2* gene was not critical. Recombinants vCF14 and vCF15 have the *lacZ* gene alone or with the human *IL-2* gene respectively, inserted into the middle of the vaccinia haemagglutinin gene¹³; vSC8 (ref. 14) and vCF7 (ref. 7) have the *lacZ* gene alone or with the *IL-2* gene respectively, inserted into the middle of the vaccinia thymidine kinase (*TK*) gene. Inactivation of both vaccinia genes had an appreciable attenuating effect, as judged by the longer prodromal period before death of nu/nu mice (Fig. 2*b* and *c*). In each case, however, *IL-2* expression greatly increased survival. Deaths were noted with vCF7 (one animal at day 18) and vCF15 (one animal at day 18 and one at day 19) at a time when all animals receiving vSC8 and vCF14 were already dead. Necropsy of the animal receiving vCF7 yielded a large amount of vaccinia in liver and spleen ($>1 \times 10^6$ PFU per 100 mg tissue). These virus plaques did not give a blue colour in the presence of Xgal (ref. 14), indicating the absence of β -galactosidase, and did not produce detectable *IL-2* bioactivity; in addition, this virus did not grow on TK⁻ cells in the presence of 5-bromodeoxyuridine, indicating that it had a TK⁺ phenotype. Therefore we believe that the inoculum given to this mouse was contaminated with wild-type virus. Necropsy of an animal receiving vCF14 yielded a small amount of virus in liver and spleen ($<2 \times 10^2$ PFU per 100 mg tissue). This virus gave blue plaques in the presence of Xgal and retained the capacity to express *IL-2* *in vitro*; the exact cause of death in this animal is under investigation.

The virulence of vaccinia recombinants in normal BALB/cByJ mice was determined by an intracranial LD₅₀ assay². It was previously reported that TK⁻ vaccinia virus recombinants were

greatly attenuated in normal mice when injected either intraperitoneally or intracranially². We confirmed the latter result by finding a 10^4 times increase in the intracranial LD₅₀ of vSC8 compared to wild-type virus ($3.2 \pm 0.6 \times 10^5$ PFU versus $4.0 \pm 1.0 \times 10^1$; $P < 0.001$, two-tailed *Z*-test¹⁵). In addition, we found that inactivation of the vaccinia haemagglutinin gene of vCF14 greatly increased the LD₅₀ ($1.4 \pm 0.3 \times 10^5$, $P < 0.001$), although only by half as much as vSC8. Inactivation of the *HindIII* *C* gene of vTFCLZ-1 or vCF22 increased the LD₅₀ only slightly ($3.0 \pm 0.5 \times 10^2$ for vTFCLZ-1, $P < 0.01$; $2.0 \pm 0.4 \times 10^3$ for vCF22, $P < 0.001$). A significant additional attenuating effect of *IL-2* expression was obtained when insertion was into the *HindIII* *C* gene (LD₅₀ for vCF13 $8.0 \pm 0.0 \times 10^5$, $P < 0.001$, compared to vTFCLZ-1 or vCF22), but not when inserted into either the *TK* (LD₅₀ for vCF7 $1.0 \pm 0.2 \times 10^6$, $P > 0.05$, compared to vSC8) or vaccinia haemagglutinin gene (LD₅₀ for vCF15 $5.6 \pm 0.9 \times 10^5$, $P > 0.05$, compared to vCF14).

Reduction in the size of pox lesions produced by TK⁻ vaccinia recombinants compared to wild-type virus injected intradermally in rabbits has been reported¹⁶. We confirmed these results with vSC8, which required 10^6 times more virus to produce the same size of skin lesions as wild-type WR in normal New Zealand white rabbits. Inactivation of the vaccinia haemagglutinin also attenuated pox formation in rabbits: vCF14 required 10^4 times more virus to produce the same size of skin lesion as WR. Insertion of *lacZ* into the *Bam*HI deletion in vTFCLZ-1 did not appreciably reduce pox size compared to WR. Skin lesion formation by vaccinia/*IL-2* recombinants showed the same degree of attenuation as seen in intracranial LD₅₀ assays: vCF13 required about 10^4 times more virus to produce the same size of skin lesion as WR or vTFCLZ-1. vCF15 and vCF7 produced lesions comparable to their respective controls (vCF14 and vSC8) at all doses (data not shown).

We were uncertain whether *IL-2* expression would enhance or diminish IgG production in normal mice. The former could occur because of immune stimulation, whereas the latter might occur if the virus were rapidly cleared. As these results would

Table 1 Antibody titres to vaccinia virus and influenza in mice vaccinated with recombinant vaccinia vectors expressing influenza antigens with β -gal, or β -gal plus human IL-2

Dose/route	Anti-influenza antibody titres		Anti-vaccinia antibody titres		
	vCF20 (HA/LacZ)	vCF19 (HA/IL-2)	vCF20 (HA/LacZ)	vCF19 (HA/IL-2)	
10 ⁷ i.v.	11.2±0.2	12.4±0.9	15.2±0.4	15.2±0.4	BALB/c
10 ⁵ i.v.	8.4±0.7	9.8±0.3	9.6±1.3	11.4±0.6	
10 ² i.v.	7.4±0.2	6.8±0.4	7.8±1.1	8.2±1.0	
10 ² i.d.	3.1±1.6	2.4±0.9	7.4±0.2	7.8±0.2	
10 ⁷ i.d.	5.6±0.5	4.6±1.2	5.2±1.3	5.4±1.2	B10.A(5R)
10 ⁵ i.v.	10.2±0.8	10.6±0.4	12.8±0.6	12.4±0.5	
10 ² i.v.	9.6±0.2	9.8±1.1	6.6±0.2	7.8±0.5	
10 ² i.d.	0.0±0.0	2.8±1.0*	0.0±0.0	1.4±1.0	
10 ² i.v.	0.0±0.0	2.6±2.6	0.0±0.0	0.2±0.2	
Dose/route	vCF12 (NP/LacZ)	vCF10 (NP/IL-2)	vCF12 (NP/LacZ)	vCF10 (NP/IL-2)	
	(NP/LacZ)	(NP/IL-2)	(NP/LacZ)	(NP/IL-2)	
10 ⁷ i.v.	8.5±0.7	7.4±0.9	14.4±0.6	13.6±1.1	BALB/c
10 ⁷ i.d.	8.2±2.9	9.6±1.7	13.6±0.6	14.5±1.7	
10 ⁵ i.v.	7.8±1.6	7.4±1.5	9.0±0.7	9.2±1.9	
10 ⁷ i.v.	2.4±2.9	7.8±1.1*	9.4±1.1	10.6±3.6	B10.A(5R)
10 ⁷ i.d.	7.2±1.0	9.2±1.0	9.4±0.5	9.2±0.4	
10 ⁵ i.v.	1.5±1.5	7.0±0.6*	1.1±1.0	6.2±1.7*	
10 ⁵ i.d.	7.2±1.7	7.4±0.9	6.6±4.2	8.0±2.5	

The virus vector, size of inoculum, route of vaccination and mouse strain are indicated. 6-week-old male BALB/c or B10.A(5R) mice were injected in the tail vein with 0.2 ml purified virus diluted in PBS, or through tail scratch with 0.03 ml. All animals were housed in adjacent cages in the same room, and matched groups of animals were injected simultaneously. Animals were bled on day 0 and day 28 after vaccination and serum was used in ELISA. ELISA results are shown, with numbers representing the mean number of positive serial twofold dilutions, starting at a serum concentration of 1:20, $\pm$ the standard error for groups of five mice. Pre-immune sera in BALB/c mice had mean anti-vaccinia titres of <1:80; pre-immune titres for all other antigens in BALB/c and B10.A(5R) mice were <1:20. Specificity of anti-influenza antibodies was demonstrated by <1:20 reactivity of immune sera in a Sendai virus ELISA (data not shown).

Methods. Double recombinants were constructed by transfecting CV-1 cells infected with vaccinia recombinant v64 (ref. 22) expressing the influenza A/PR/8/34 haemagglutinin with pCF11 or pCF13 (see Fig. 2), yielding recombinant viruses vCF20 and vCF19 respectively. CV-1 cells infected with vaccinia recombinant v69 (ref. 23) expressing the influenza A/PR/8/34 nucleoprotein were transfected with pCF11 or pCF13, yielding recombinants vCF12 and vCF10 respectively. Recombinant vectors were characterized, purified, and titred on BSC-1 cells as described^{14,20}. Vectors vCF19 and vCF10 expressed active IL-2 to the same extent as vCF13 (see text). ELISAs for vaccinia virus and influenza A were modified from previously reported protocols^{24,25}. For vaccinia virus, microtitre plates were coated overnight at 4°C with 0.2 µg per well purified vaccinia virus (WR), diluted in bicarbonate buffer (15mM Na₂CO₃, 35mM NaHCO₃, pH 9.6). Plates were washed extensively with PBS-0.5% Tween 20 (Sigma), and unbound sites blocked with 5% ovalbumin (Grade V, Sigma) in PBS for 6 h at room temperature. Plates were washed and serial twofold dilutions of serum made across the plate, diluted in PBS/Tween 20. After overnight incubation at 4°C, plates were washed and incubated with a 1:4,000 dilution of peroxidase-conjugated rabbit anti-mouse IgG (Southern Biotechnology). After overnight incubation, plates were washed and reacted with peroxidase substrate (ABTS, Kirkegaard and Perry) for 1 h at 20°C. Optical density of wells was read at 405 nm. ELISAs for influenza were similar: wells were coated with 20 haemagglutination units per well of purified influenza A/PR/8/34, solubilized with 0.1% SDS, and dried on to wells overnight for anti-NP titres, or 8 haemagglutination units per well purified influenza A/PR/8/34 in bicarbonate buffer incubated overnight at 4°C for anti-HA titres.

* Pairs significantly different, $P < 0.01$, Student's paired t -test.

be particularly important if IL-2 recombinants were to be used as vaccines, we constructed double recombinant vaccinia viruses containing the IL-2 or lacZ gene in the HindIII C locus and either the influenza haemagglutinin (HA) or nucleoprotein (NP) gene in the TK locus. Inoculations were given to groups of BALB/c (K^dIa^dD^d) or B10.A(5R) mice (K^bIa^bD^d) over a range of dilutions either by the intravenous (i.v.) or intradermal (i.d.) routes. Antibody titres were dependent on the size and route of inoculum as well as the mouse strain (Table 1). BALB/c mice showed higher antibody titres than B10.A(5R) to both vaccinia and influenza antigens; influenza HA titres were higher than NP titres, and with low-virus inocula the i.d. route was generally superior to the i.v. route. The IL-2-expressing and non-expressing recombinants gave similar antibody titres in BALB/c mice. In B10.A(5R) mice, the IL-2 recombinants produced sig-

nificantly higher influenza antibody titres with all i.v. doses of the NP recombinant and a low i.v. dose (10³ PFU) of the HA recombinant. In both strains of mice, protection against a lethal intranasal challenge with influenza A/PR/8/34 (10 mice per group: 100 LD₅₀ for influenza HA recombinants; 10 LD₅₀ for NP) was complete with vaccinia/HA recombinant doses as low as 10⁵ PFU, partial (40%) at 10³ PFU, and absent at 10² PFU. Coexpression of IL-2 did not influence protection at any dose ($p > 0.05$, Fisher's Exact Test). Vaccinia/NP recombinants provided partial protection against mortality (35–55%) in BALB/c mice compared to mice vaccinated with equivalent doses of vSC8, as previously described^{17,18}. Vaccinia/NP recombinants provided no specific protection against influenza mortality in B10.A(5R) mice. In all cases, protection was not enhanced by IL-2 expression.

These studies demonstrate that attenuation of vaccinia virus virulence can be achieved by expression of a lymphokine such as human IL-2, as well as by deletion of certain vaccinia virus genes. IL-2 expression was particularly effective for immunodeficient athymic nude mice. Although these mice have few T lymphocytes with which to combat infection, IL-2 expressing vaccinia virus was rapidly cleared following systemic administration. There is no evidence that clearance was achieved by enhanced anti-vaccinia antibodies, cytotoxic lymphocytes or natural killer cell activity. The possibility is being investigated that IL-2 secretion by vaccinia-infected cells induces significant local production of interferon or stimulation of lymphokine-activated killer cells, as has been shown when IL-2 is administered systemically¹⁹. Attenuation as a result of IL-2 expression also occurred in normal mice inoculated intracranially with vaccinia virus, although this effect was apparently masked when the recombinants were already highly attenuated due to insertional mutagenesis. Studies with double recombinant vaccinia viruses containing influenza virus genes indicated that IL-2 expression did not result in loss of immunogenicity in normal mice. In fact, antibody production was enhanced under certain conditions, such as with a poor responder strain of mouse, a weak immunogen, or unfavourable route of inoculation. These results are consistent with the only published report of IL-2 enhancement of humoral immunity¹¹.

Lymphokine co-expression is a novel method for altering the virulence of viruses without adversely affecting immunogenicity. This represents an active approach to attenuation that differs fundamentally from procedures involving gene inactivation. Although it has been shown here to be effective with vaccinia virus, IL-2 or other lymphokines could have beneficial effects with other candidate live virus vaccines.

We are grateful to Norman Cooper for assistance with tissue culture and preparation of purified vaccinia stocks, and to Drs S. Arya, S. Broyles, R. M. L. Buller, P. Earl, T. Fuerst, G. Kotwal, S. Rosenberg, S. Schwartz, A. Vargosko, F. Wong-Staal, J. Yewdell and T. Yilma for supplying essential materials and advice. A.H. was supported by a fellowship from the Schweizerischer Nationalfond.

Received 17 August; accepted 24 September 1987.

1. Moss, B. & Flexner, C. A. *Rev. Immun.* **5**, 305–324 (1987).
2. Buller, R. M. L., Smith, G. L., Cremer, K., Notkins, A. L. & Moss, B. *Nature* **317**, 813–815 (1985).
3. Redfield, R. R. *et al. N. Engl. J. Med.* **316**, 673–676 (1987).
4. Clark, S. C. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 2543–2547 (1984).
5. Rosenberg, S. A. *et al. Science* **223**, 1412–1415 (1984).
6. Taniguchi, T. *et al. Nature* **302**, 305–310 (1983).
7. Flexner, C., Hügin, A. & Moss, B. in *Vaccines 87: Modern Approach to New Vaccines including Prevention of AIDS* (eds Chanock, R. M., Lerner, R. A., Brown, F. & Ginsberg, H.) 380–383 (Cold Spring Harbor, New York, 1987).
8. Rosenberg, S. A. & Lotze, M. T. *et al. Rev. Immun.* **4**, 681–709 (1986).
9. Rouse, B. T., Miller, L. S., Turtinen, T. & Moore, R. N. *J. Immun.* **134**, 926–930 (1985).
10. Weinberg, A., Rasmussen, L. & Merigan, T. C. *J. infect. Dis.* **154**, 134–140 (1986).
11. Kawamura, H., Rosenberg, S. A. & Berzofsky, J. A. *J. exp. Med.* **162**, 381–386 (1985).
12. Mackett, M., Smith, G. L. & Moss, B. *Proc. natn. Acad. Sci. U.S.A.* **79**, 7415–7419 (1982).
13. Shida, H. *Virology* **150**, 451–462 (1986).
14. Chakrabarti, S., Brechling, K. & Moss, B. *Molec. cell. Biol.* **5**, 3403–3409 (1985).
15. Miller, R. G. *Biometrika* **60**, 535–542 (1973).

16. Buller, R. M. L. & Moss, B. in *Vaccinia Viruses as Vectors for Vaccine Antigens* (ed. Quinnan, G. V.) 37-47 (Elsevier, New York, 1985).
17. Andrew, M. E., Coupar, B. E. H., Ada, G. L. & Boyle, D. B. *Microbial Pathogenesis* 1, 443-452 (1986).
18. Wraith, D. C., Vessey, A. E. & Askonas, B. A. *J. gen. Virol.* 68, 433-440 (1987).
19. Weinberg, A., Konrad, M. & Merigan, T. C. *J. Virol.* 61, 2120-2127 (1987).
20. Mackett, M., Smith, G. L. & Moss, B. in *DNA Cloning*, Vol. II (ed. Glover, D. M.) 191-212 (IRL, Oxford, 1985).
21. Nakano, J. H. in *Diagnostic Procedures for Viral, Rickettsial, and Chlamydial Infections*, 5th edn, (eds Lennette, E. H. & Schmidt, N. J.) 296-299, (American Public Health Association, Washington DC, 1979).
22. Binnink, J. R., Yewdell, J. W., Smith, G. L. & Moss, B. *J. Virol.* 57, 786-791 (1986).
23. Yewdell, J. W., Binnink, J. R., Smith, G. L. & Moss, B. *Proc. natn. Acad. Sci. U.S.A.* 82, 1785-1789 (1985).
24. Buller, R. M. L., Bhatt, P. N. & Wallace, G. D. *J. clin. Microbiol.* 18, 1220-1225 (1983).
25. Murphy, B. R. *et al. J. clin. Microbiol.* 13, 554-560 (1981).

Major histocompatibility complex-linked specificity of $\gamma\delta$ receptor-bearing T lymphocytes

Louis A. Matis*, Randy Cron†
& Jeffrey A. Bluestone†‡

* Molecular Immunology Laboratory, Division of Biochemistry and Biophysics, Center for Drugs and Biologics, Food and Drug Administration, Bethesda, Maryland 20892, USA

† Transplantation Biology Section, Immunology Branch, National Cancer Institute, NIH, Bethesda, Maryland 20892, USA

Several recent studies have identified a distinct subset of CD3(T3)⁺CD4⁻CD8⁻ T lymphocytes that express a CD3-associated heterodimer made up of the protein encoded by the T-cell receptor (TCR) γ -gene and a second glycoprotein termed TCR δ (refs 1-4). TCR $\gamma\delta$ is expressed on CD3⁺ thymocytes during fetal ontogeny before the appearance of TCR alpha-beta ($\alpha\beta$) (refs 5-7), on CD3⁺CD4⁻CD8⁻ adult thymocytes^{4,6}, and on a subset (1-10%) of CD3⁺ cells in adult peripheral lymphoid organs and the peripheral blood^{1,8,9}. TCR $\gamma\delta$ -expressing T cells probably represent a distinct mature T-cell lineage with the capacity to proliferate in response to receptor-mediated signals^{2,6}, and to display non-major histocompatibility complex (MHC)-restricted cytotoxicity^{2,10-13}. Critical to understanding the function of this T-cell subset is the identification of the ligand(s) recognized by TCR $\gamma\delta$. Here we describe an alloreactive CD3⁺CD4⁻CD8⁻TCR $\gamma\delta$ -expressing, TCR $\alpha\beta$ -negative, T-cell line that manifests MHC-linked recognition specificity for both proliferation and cytotoxicity. Our results suggest that T cells expressing TCR $\gamma\delta$ are capable of self-non-self MHC discrimination and that they can undergo MHC-influenced selection during differentiation like TCR $\alpha\beta$ -expressing T cells.

As an approach to identifying a potential MHC-linked specificity of TCR $\gamma\delta$ -expressing T cells, an alloreactive T-cell line was established from the peripheral lymph nodes of immunized BALB/c (H-2^d) nu/nu mice by stimulation every 7-10 days¹⁴ with irradiated allogeneic low-density B10.BR (H-2^k) spleen cells (APC). Four weeks after the start of culture, flow cytometric analysis showed that the cell population was 100% CD3⁺ (Fig. 1a). Among the CD3⁺ cells, 74% were CD4⁻CD8⁻ (Fig. 1a) and 26% were CD8⁺ (Fig. 1b). TCR analysis was also performed using a monoclonal antibody F23.1 (ref. 15) which is specific for variable (V)-region proteins of the V β 8 gene subfamily¹⁶. At four weeks, V β 8⁺ cells constituted about 5% of the total population (Fig. 1b). All the V β 8⁺ cells (implying all the TCR $\alpha\beta$ -expressing cells) were encompassed in the CD8⁺ population (Fig. 1b). After further culture, the number of CD8⁺ cells (as well as V β 8⁺ cells) progressively diminished, so that by week 8 (Fig. 1) and thereafter (data not shown), all the cells were CD3⁺CD4⁻CD8⁻ V β 8⁻ (Fig. 1).

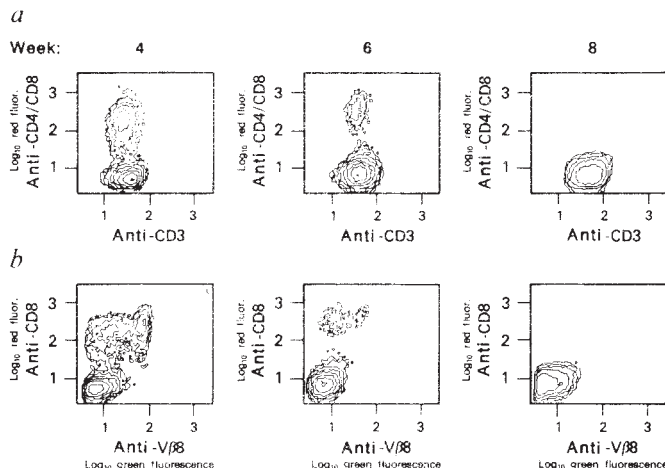


Fig. 1 Surface phenotypic characterization of the BALB/c nu/nu-derived proliferative T-cell line. The two-colour data of the flow cytometry analysis are displayed as contour diagrams with logarithmically increasing intensities of green (FITC) fluorescence plotted on the x-axis against logarithmically increasing intensities of red (Texas-Red) fluorescence on the y-axis. Surface phenotypic characterization of the cell line is shown after 4, 6 and 8 weeks of *in vitro* culture. *a*, Anti-CD3 staining on the x-axis versus anti-CD4/CD8 staining on the y-axis. *b*, anti-V β 8 staining on the x-axis versus anti-CD8 staining on the y-axis. V β 8 is expressed on about 20% of splenic TCR $\alpha\beta$ -bearing T cells in the BALB/c mouse²⁹. **Methods.** 20-week-old BALB/c nu/nu mice were immunized in the hind footpads with 10⁷ irradiated (3,300R) B10.BR spleen cells emulsified in complete Freund's adjuvant (type H37RA, Difco). One week later, draining lymph nodes were collected and single cell suspensions prepared. Bovine serum albumin (BSA)-gradient-enriched low density B10.BR spleen cells were used as antigen-presenting cells (APC). To prepare APC, B10.BR spleen cells were suspended in 5 ml of 35% BSA (Miles); 5 ml 11% BSA were overlaid, and the cells were spun for 30 min at 12,000 r.p.m. at 4°C. Cells at the interface were collected, washed once in RPMI 1640, and resuspended in complete medium consisting of a 1:1 mixture of RPMI and EHAA, 10% fetal bovine serum, 15 mM HEPES buffer, 4 mM glutamine, 100 μ g ml⁻¹ gentamycin, 100 U ml⁻¹ penicillin, and 2.5 $\times$ 10⁻⁵ M indomethacin (Sigma). The T cells were cultured at 5 $\times$ 10⁶ per well in flat-bottom 24-well tissue culture plates in the presence of 10⁶ irradiated (2,000R) APC. The cells were collected after 10 d, the viable cells purified over Ficoll gradients (Pharmacia), washed, and 2 $\times$ 10⁶ T cells were cultured with 2 $\times$ 10⁶ APC. On the third and subsequent stimulations, the responding T cells were collected 2 d following activation with allogeneic APC, and recultured at 2 $\times$ 10⁵ cells per well in the presence of 100 U per ml of recombinant human interleukin-2 (IL-2) (Lot LP315, Cetus)²⁵. Two days before each restimulation, and before running proliferation and cytotoxicity assays, cells were recultured in complete medium without exogenous IL-2. Multi-colour flow cytometry analysis was performed as described⁶.

To define the CD3-associated receptor complex, surface radioiodinated cells were immunoprecipitated with a monoclonal antibody specific for the murine CD3- ϵ protein (145-2C11) (ref. 17) and with an antiserum raised against murine TCR γ (ref. 4) (Fig. 2). The antibody against CD3 immunoprecipitated a protein of relative molecular mass (M_r) 80,000 (80K) under non-reducing conditions (Fig. 2a, lane 1), which reduced to 35K and 45K proteins (Fig. 2a, lane 2), a result consistent with the 80K protein being a disulphide-linked dimer. The control hamster-derived monoclonal antibody 145-10D8 (ref. 6) failed specifically to precipitate any bands (Fig. 2a, lane 3). The antiserum against TCR γ also immunoprecipitated 35K and 45K proteins (Fig. 2a, lane 4), identical in size to the previously described TCR γ - and δ -proteins precipitated from lysates of surface-labelled murine CD3⁺CD4⁻CD8⁻ fetal⁵ and adult⁴ thymocytes. Immunoprecipitation of both 35K and 45K bands by the anti-TCR γ serum was inhibited by excess immunizing γ peptide (Fig. 2a, lane 5). Both bands were distinct from

‡ Present address: Department of Pathology, University of Chicago, Chicago, Illinois 60637, USA.

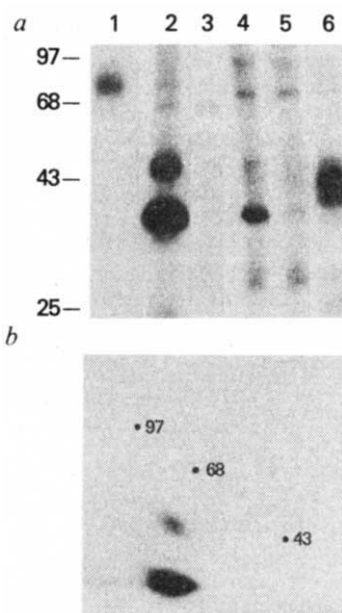


Fig. 2 Biochemical characterization of the CD3-associated proteins on the CD3⁺CD4⁺CD8⁺BALB/c nu/nu anti-B10.BR alloreactive T cells. Analyses were performed on 12 week (or later) cultured cells. Specific immunoprecipitations used the anti-CD3 monoclonal antibody 145-2C11 (ref. 17) and an affinity-purified anti-TCR γ antiserum¹ on cells from the BALB/c nu/nu T-cell line (Fig. 2a, lanes 1–5; Fig. 2b) or on ACK-lysed murine spleen cells (Fig. 2a, lane 6). *a*, Lane 1, anti-CD3, non-reducing conditions; lane 2, anti-CD3, reducing conditions; lane 3, specificity control hamster-derived monoclonal antibody 145-10D8, reducing conditions; lane 4, anti-TCR γ antiserum, reducing conditions; lane 5, anti-TCR γ , 100 μ g immunizing γ peptide, reducing conditions; lane 6, anti-CD3, normal spleen cells, reducing conditions. *b*, Anti-CD3 immunoprecipitates of the BALB/c nu/nu T-cell line run under non-reducing conditions in the first dimension and then under reducing conditions in the second dimension. M_r of standard proteins are shown (K).

Methods. The T cells were surface radioiodinated as described⁴ and then lysed in 1% digitonin buffer⁵ (for anti-CD3 immunoprecipitations) or 0.5% NP-40 (for anti-TCR γ precipitations)⁵. Lysates were precleared overnight at 4 °C with rabbit serum and immunoprecipitated as described⁵. Immune complexes were collected with protein A agarose beads, washed 4 times in lysing buffer, dissociated by heating for 2 min at 90 °C in 2% SDS, and run under reducing conditions (5% 2-mercaptoethanol) or non-reducing conditions on 10% SDS-PAGE gels for 18 h at 6 mA. The gels were dried and labelled proteins visualized using Kodak XAR-2 radiographic film at –70 °C with enhancing screens. For inhibition of anti-TCR γ immunoprecipitation, 100 μ g γ peptide⁴ was added during the immunoprecipitation. For the two-dimensional gel analysis, samples were eluted in non-reducing sample buffer and run on a 10% SDS-PAGE tube gel for 14 h at 0.25 mA per tube. The tube gels were then incubated in 5 ml of reducing sample buffer for 2 h at room temperature, adhered to a 10% SDS-PAGE slab gel with 1% agarose containing 5% 2-mercaptoethanol in sample buffer, and run under reducing conditions for 18 h at 6 mA. The gels were dried and visualized as above.

the TCR $\alpha\beta$ proteins (43–37K) observed after anti-CD3 immunoprecipitation of normal splenic T cells (Fig. 2a, lane 6). Thus, this T-cell line expressed exclusively the CD3-associated TCR $\gamma\delta$ heterodimer.

The disulphide-linked heterodimeric nature of the 35K and 45K proteins was confirmed using two-dimensional nonreducing/reducing gel electrophoresis of anti-CD3 immunoprecipitates^{4,5}. Both the CD3-associated 35K and 45K proteins migrated below the diagonal (Fig. 2b), indicating that they migrate as a single 80–90K molecule in the first (non-reduced) dimension. N-glycosidase F analysis shows that all the 35K TCR γ proteins

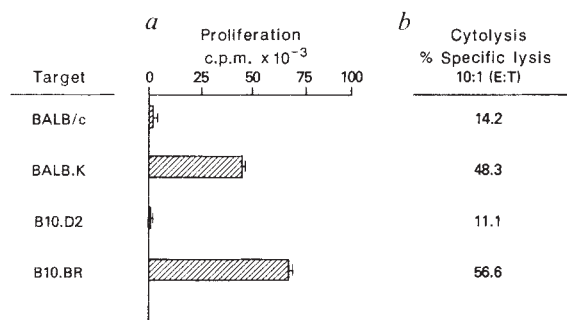


Fig. 3 Demonstration of the MHC-linked specificity of proliferation and cytotoxicity of the BALB/c nu/nu anti-B10.BR T cell line. *a*, Proliferation of BALB/c nu/nu T cells in response to various H-2 congenic stimulator cells (BALB/c, H-2^d; BALB.K, H-2^k; B10.D2, H-2^d; B10.BR, H-2^k) is represented as the arithmetic mean ($\pm$ s.e.m.) of triplicate cultures. *b*, Cytolytic specificity assayed on the same four target cell populations. The % specific lysis was determined as follows: % specific lysis = [(exp. c.p.m. – spont. c.p.m.)/(total c.p.m. – spont. c.p.m.)] $\times 100$. “Exp. c.p.m.” represents the release of ⁵¹Cr measured in the experiment from target cells incubated with effector cells. “Spont. c.p.m.” represents the spontaneous release of ⁵¹Cr from targets incubated in the absence of effector cells, and was always <25% of total c.p.m. The total c.p.m. represents the amount of ⁵¹Cr released by the target cells after incubation in a solution of 0.05M HCl. The values shown are the means of triplicate determinations. Standard errors, <10%. **Methods.** *a*, 2×10^6 responding T cells were cultured in complete medium with 2×10^6 irradiated (2,000R) APC in 24-well flat-bottom plates in a total volume of 2 ml per well. After 24 h, the cells were resuspended, and 0.2 ml per well was added to individual wells of a 96-well flat-bottom microtitre plate. Each well was then pulsed with 1 μ Ci of tritiated thymidine for 12 h. Stimulator cells were prepared as described in the legend to Fig. 1. *b*, ⁵¹Cr release assay. Target cells were radiolabelled by incubating $2-4 \times 10^6$ concanavalin A (Miles-Yeda)-induced splenic blasts with 300 μ Ci Na⁵¹CrO₃ in 0.3 ml FCS for 1 h. Radiolabelled target cells were washed three times, counted and resuspended in medium at 5×10^5 cells ml⁻¹. The T cells were collected and resuspended at 5×10^6 cells ml⁻¹ and then dispensed into 96-well U-bottom microculture plates (Costar). The radiolabelled target cells were added to each well in 100 μ l medium. After incubating for 4–5 h at 37 °C, the culture supernatants were collected with the Titertek supernatant collection system (Skatron).

in the line are glycosylated (R.C., L.A.M. and J.A.B., unpublished data), which suggests that they are predominantly the products of V γ -C γ 1 rearranged genes because V γ -C γ 1 proteins possess N-linked carbohydrate acceptor sites, but V γ 1.2–C γ 2 proteins do not^{4,5,19}.

The antigen specificity of the BALB/c nu/nu T-cell line was examined in both proliferation and cytotoxicity assays. Vigorous proliferation was observed in response to H-2 congenic BALB.K (H-2^k), as well as to B10.BR (H-2^k) APC, whereas neither syngeneic BALB/c (H-2^d) nor B10.D2 (H-2^d) APC elicited significant proliferative responses (Fig. 3a). The T cells were also cytotoxic with a specificity identical to the proliferative response (Fig. 3b). The specificity of MHC recognition was further explored in cytotoxicity assays using target cell populations from various H-2 recombinant strains of mice (Fig. 4). For example, the data in Fig. 4a show significant specific lysis of B10.BR (K^kI^kD^k) and C3H.OL (K^dI^dD^k) target cells, but not B10.A (K^kI^kD^d) or B10.D2 (K^dI^dD^d) target cells. Moreover, the cells lyse an H-2^k tumour cell line (R1.1), but not a β 2-microglobulin negative variant (R1.E) of the same tumour cell line²⁰ (J.A.B. and L.A.M., unpublished data). These data are consistent with the recognition of class I MHC molecules encoded within either the H-2D or the adjacent Qa or TLa gene regions²¹.

Further analysis of other MHC-congenic target cells demonstrated a reproducible enhanced lysis of allogeneic B10 (K^bI^bD^b) target cells (Fig. 4b). Effective lysis of B10.GD (K^dI^dD^b), but

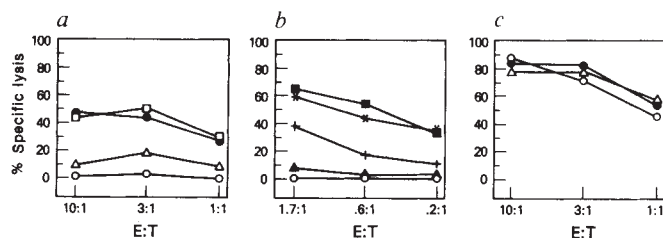


Fig. 4 Analysis of the MHC-linked cytolytic specificity of the BALB/c nu/nu anti-B10.BR TCR $\gamma\delta$ -expressing alloreactive T cell line. *a* and *b*, Specificity of resting (prestimulation) T cells. *c*, Specificity of IL-2-stimulated T cells. The target cell populations were as follows: *a*, $\square$ —, B10.D2; $\square$ —, C3H.OL; $\triangle$ —, B10.A; $\bullet$ —, B10.BR. *b*, $\square$ —, B10.D2; $\blacksquare$ —, B10.GD; $\triangle$ —, B10.YBR; $\ast$ —, B10; --- , B10.Q. *c*, $\square$ —, B10.D2; $\bullet$ —, B10.BR; $\triangle$ —, B10.A.

Methods. The cytolytic (^{51}Cr release) assays were performed as described in the legend to Fig. 3. To examine the effect of IL-2 on cytolytic specificity (*c*), the T cells which had been stimulated with allogeneic B10.BR APC 2d previously were then cultured in the presence of 100 U ml^{-1} recombinant human IL-2 (ref. 25) for 3 d, collected, and immediately assayed for cytolytic specificity.

not B10.YBR (K^bI^D^d), target cells also maps this H-2^b -specific cross-reactive cytolytic activity to the H-2D , Qa , or TLa regions of the H-2^b locus. The T cells also lyse B10.Q (H-2^q) target cells (Fig. 4*b*). Thus, these TCR $\gamma\delta$ T cells display a broad pattern of cross-reactivity for allogeneic target cells. This degree of cross-reactivity is in contrast to that expected from class I MHC-specific TCR $\alpha\beta$ alloreactive T cells²², and suggests the recognition of relatively non-polymorphic antigenic determinants. Recently we have derived several T-cell clones from the line; all of these are $\text{CD3}^+ \text{CD4}^- \text{CD8}^-$, express the $\gamma\delta$ -receptor heterodimer, and manifest MHC-linked cytolytic specificity identical to that of the line, including the cross reactivities for H-2^b and H-2^q target cells (data not shown).

The MHC-linked specificity of our TCR $\gamma\delta$ -expressing cell line contrasted with previously reported examples of non-MHC-specific cytotoxic activity of TCR $\gamma\delta$ -expressing cells^{2,10,11}, but these cell lines were established in the continuous presence of growth factors, including IL-2 (refs 2, 10–12). Because our T-cell line was assayed for specificity after culture in the absence of growth factor for at least 2 days, we reasoned that the addition of exogenous IL-2 might induce nonspecific lytic activity, as described previously for $\text{CD4}^- \text{CD8}^+$ class I MHC-specific TCR $\alpha\beta$ T cells^{23,24}, and this was found to be the case. The BALB/c nu/nu T cells assayed immediately after 3 days of continuous culture in the presence of 100 U per ml of recombinant human interleukin-2 (ref. 25) clearly mediated non-MHC-specific cytolytic activity, as shown by the lysis of self-MHC (H-2^d)-expressing B10.D2 target cells (Fig. 4*c*).

Thus the data presented here establish that alloreactive TCR $\gamma\delta$ -expressing T cells with proliferative and cytolytic function can recognize an MHC-linked antigen, probably expressed on class I MHC molecules. The derivation of such cells from the peripheral lymph nodes of nu/nu mice demonstrates that they can differentiate through an extrathymic pathway. A similar extrathymic developmental pathway has also been described for class I MHC-restricted $\text{CD4}^+ \text{CD8}^+$ TCR $\alpha\beta$ -bearing cytotoxic T cells²⁶, but not for $\text{CD4}^+ \text{CD8}^-$ class II MHC-restricted T cells²⁷. This parallel between class I MHC-specific TCR $\alpha\beta$ - and TCR $\gamma\delta$ -expressing T cells suggests that T cells might generally mature extrathymically for the recognition of class I but not class II MHC molecules.

These results do not preclude the possibility that we have selected a small subset of MHC-specific TCR $\gamma\delta$ -bearing T cells by virtue of the activation with allogeneic APC, and that many TCR $\gamma\delta$ T cells recognize non-MHC encoded antigens. But the immunization and the *in vitro* stimulation with B10.BR APC,

which bear disparate non-MHC as well as MHC antigens, should have enhanced the likelihood of deriving $\gamma\delta$ T cells specific for non-MHC encoded determinants. Also, it could be relevant that in-frame TCR γ messenger RNA is induced after allogeneic (but not lectin) stimulation of murine splenic T cell populations²⁸. Finally, it remains to be determined whether TCR $\gamma\delta$ -expressing T cells can recognize non-MHC-encoded foreign antigens in a self-MHC-restricted fashion.

We thank Drs Richard Hodes, Drew Pardoll and Larry Samelson for critical review of the manuscript, Dr John Coligan for the anti-TCR γ antiserum, Ms Michele Foo for technical assistance, and Ms Ellen Kirshbaum for preparation of the manuscript. R.C. is a Research Scholar of the Howard Hughes Medical Institute.

Received 9 July; accepted 17 September 1987.

1. Brenner, M. B. *et al.* *Nature* **322**, 145–149 (1986).
2. Bank, I. *et al.* *Nature* **322**, 179–181 (1986).
3. Weiss, A., Newton, M. & Crommie, D. *Proc. natn. Acad. Sci. U.S.A.* **83**, 6998–7002 (1986).
4. Lew, A. M. *et al.* *Science* **234**, 1401–1405 (1986).
5. Pardoll, D. M. *et al.* *Nature* **326**, 79–81 (1987).
6. Bluestone, J. A., Pardoll, D., Sharrow, S. O. & Fowlkes, B. J. *Nature* **326**, 82–84 (1987).
7. Nakanishi, N., Maeda, K., Ito, K., Heller, M. & Tonegawa, S. *Nature* **325**, 770–773 (1987).
8. Lanier, L. L. & Weiss, A. *Nature* **324**, 268–270 (1986).
9. Lanier, L. L. *et al.* *J. exp. Med.* **165**, 1076–1094 (1987).
10. Brenner, M. B. *et al.* *Nature* **325**, 689–694 (1987).
11. Borst, J. *et al.* *Nature* **325**, 683–688 (1987).
12. Moingeon, P. *et al.* *Nature* **325**, 723–726 (1987).
13. Reinherz, E. *Nature* **325**, 660–663 (1987).
14. Matis, L. A. *et al.* *J. Immun.* **135**, 803–713 (1985).
15. Staerz, U. D., Rammensee, H. G., Benedetto, J. D. & Bevan, M. J. *J. Immun.* **134**, 3994–4000 (1985).
16. Sim, G. K. & Augustin, A. A. *Cell* **42**, 89–92 (1985).
17. Leo, O., Foo, M., Sachs, D. H., Samelson, L. E. & Bluestone, J. A. *Proc. natn. Acad. Sci. U.S.A.* **84**, 1374–1378 (1987).
18. Saito, H. *et al.* *Nature* **309**, 757–762 (1984).
19. Garman, R. D., Doherty, P. J. & Raulet, D. H. *Cell* **45**, 733–742 (1986).
20. Hyman, R. & Stallings, V. *Immunogenetics* **3**, 75–84 (1976).
21. Steinmetz, M., Stephan, O. & Lindahl, K. F. *Cell* **44**, 895–904 (1986).
22. Sherman, L. A. *Nature* **297**, 511–513 (1982).
23. Brooks, C. G. *Nature* **305**, 155–158 (1983).
24. Shortman, K. & Wilson, A. *Curr. Top. Microbiol. Immun.* **116**, 111–119 (1986).
25. Wang, A., La, S.-D. & Mark, D. *Science* **224**, 1431–1433 (1984).
26. Kruisbeek, A. M., Sharrow, S. O., Mathieson, B. J. & Singer, A. *J. Immun.* **127**, 2168–2176 (1981).
27. Kruisbeek, A. M., Davis, M. L., Matis, L. A. & Longo, D. L. *J. exp. Med.* **160**, 839–857 (1984).
28. Jones, B., Mjølness, S., Janeway, C. Jr. & Hayday, A. C. *Nature* **323**, 635–638 (1986).
29. Haskins, K. *et al.* *J. exp. Med.* **160**, 452–471 (1984).

Inhibition of mitosis by an antibody to the mitotic calcium transport system

Mathias Hafner & Christian Petzelt

Institute of Cell and Tumor Biology, German Cancer Research Center, 6900 Heidelberg, FRG

Calcium ions are important in the regulation of mitotic apparatus assembly and in the control of chromosome movement^{1–13}. Changes in intracellular free calcium concentration, $[\text{Ca}^{2+}]_i$, are achieved by an intracellular calcium-transport system which is highly conserved in different cell types^{14,15}. A membrane-bound protein of relative molecular mass (M_r) 46,000 (46K) is part of this transport system and has been implicated in the regulation of the $[\text{Ca}^{2+}]_i$ changes associated with the course of mitosis¹⁶. A monoclonal antibody against this 46K protein inhibits Ca^{2+} -uptake into isolated Ca^{2+} -sequestering membranes and specifically labels membranes associated with the mitotic apparatus of sea urchin embryos¹⁶. Here we investigate the relationship between the intracellular calcium transport system and mitosis by injection of this monoclonal antibody into living mitotic sea urchin embryos. We find that after injection the intracellular free calcium increases up to 10^{-6}M , the mitotic apparatus is rapidly destroyed and the cell is irreversibly blocked in its development.

Microinjection of the antibody against the 46K protein into mitotic sea urchin embryos of *Lytechinus pictus* results in the

Fig. 1 Effect of the microinjection of antibody to 46K protein into a two-cell sea urchin embryo at mitosis. *a*, The birefringence of the two sets of mitotic apparatus in the blastomeres immediately before the injection indicates the mitotic stage. *b*, Several minutes after the injection (indicated by the oil drop) the birefringent mitotic apparatus has disappeared and the injected cell is arrested, whereas the untreated sister-blastomere has divided. *c*, The injected blastomere remains undivided for the time of observation, but the control blastomere has developed into a blastula-like stage. *d*, SDS-PAGE showing purity of the injected antibody: lane A, IgM standard; lane B, ascites fluid; lane C, antibody to 46K protein after affinity purification (anti-IgM-agarose, Sigma). *a'*-*d'*, The result of the injection of unrelated monoclonal antibodies. No effects on the development are found. Scale bar 20 μm .

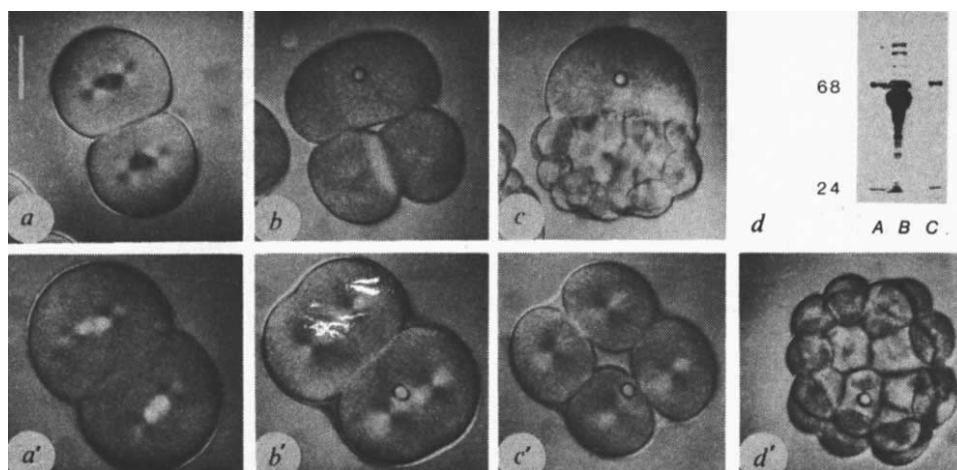
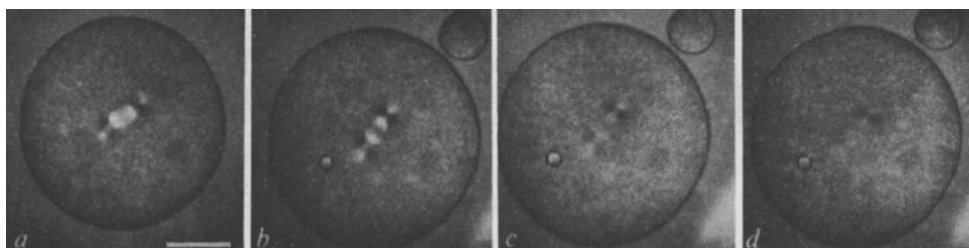


Fig. 2 Time course of the decrease of birefringence of the mitotic apparatus in a one-cell sea urchin embryo after the injection of antibody to 46K protein at mitosis. The oil drop indicates the site of injection. Pictures were taken at 1 min intervals (*b*-*c*). Note the gradient in the decrease of birefringence starting from the point of injection. Scale bar 30 μm .



arrest of mitosis (Table 1). 12-25 μl of antibody were injected, which represents 2-5% of the egg volume. Using polarized light, we monitored the effect of the injected antibody on the birefringence of the mitotic apparatus: injection of the antibody (4 mg ml^{-1}) into one of the two blastomeres at metaphase results in the complete disappearance of the birefringence of the mitotic apparatus and the subsequent developmental arrest of that cell (Fig. 1), an effect obtained only when the final intracellular concentration of the injected antibody exceeds 100 pg per cell . Up to 12 hours after injection the non-injected sister blastomere continues to progress through the normal cell cycles and eventually forms a half-blastula, whereas the injected blastomere never divides (Fig. 1*c*). Injected controls include unrelated

Table 1 Injection of the antibody against the 46K protein into mitotic *Lytechinus pictus* embryos

Solution injected	IgM (mg ml^{-1})*	No. of mitotic cells injected	Mitotic arrest
Anti-46K-protein	2-4	14	14
Denatured anti-46K-protein†	2-6	4	0
Anti-hyalin‡	2-8	5	0
Anti-histone‡	2-8	4	0
BSA	5-15‡	3	0
Mouse immunoglobulin	10-20‡	3	0
Injection buffers§	—	4	0

* Needle concentration.

† Affinity-purified monoclonal antibody.

‡ Total protein concentration.

§ 100 mM potassium aspartate, 10 mM, HEPES (pH 7.0).

|| Microinjection of affinity-purified monoclonal antibody into sea urchin embryos of *Lytechinus pictus* caused arrest of mitosis in 14 out of 14 trials. Direct pressure injection of other IgM monoclonal antibodies or serum proteins or buffer did not affect either birefringence of the mitotic apparatus or progress through mitosis and further development.

monoclonal antibodies or denatured antibody to the 46K protein at the same final intracellular concentration as the native antibody, or the buffer itself; no effect on the birefringence of the mitotic apparatus or on further cell division is found (Table 1 and Fig. 1*a'*-*d'*). After injection of the antibody to the 46K protein, the birefringence of the mitotic apparatus decreases dramatically within 60-90 seconds (Fig. 2). This effect is seen first at the point of injection and subsequently spreads over the entire mitotic apparatus (Fig. 2). If the disintegration of the mitotic apparatus and the block of mitosis is the result of interference by the injected antibody with the intracellular Ca^{2+} -transport system, it should be possible to relate the observed phenomena directly to an altered $[\text{Ca}^{2+}]_i$. Therefore, we followed $[\text{Ca}^{2+}]_i$ changes in fura2-loaded cells injected with the antibody. A natural increase in $[\text{Ca}^{2+}]_i$ occurs at the onset of mitosis (Fig. 3*a*)⁹. After the injection of the antibody into one of the two mitotic blastomeres the $[\text{Ca}^{2+}]_i$ is increased further up to 10^{-6}M within 10 minutes (Fig. 3*a*). This effect was not observed in the sister blastomere. It cannot be attributed to an unequal distribution of fura2 in both blastomeres because the dye was injected before the first cell division. Concomitantly with the increase in $[\text{Ca}^{2+}]_i$ the injected cell becomes arrested. These results indicate that interference with one of the constitutive elements of the intracellular Ca^{2+} -transport system, the 46 K protein, causes a variety of closely linked events to occur. $[\text{Ca}^{2+}]_i$ increases, the birefringence of the mitotic apparatus disappears, and mitosis is blocked. The rapid disintegration of the mitotic spindle is induced by direct injection of Ca^{2+} (refs 5, 6) or Ca-calmodulin¹⁷, and has been taken as evidence for a Ca^{2+} -lability of the mitotic apparatus *in vivo*. The rapid reappearance of an intact mitotic apparatus shows the presence of a powerful Ca^{2+} -sequestering system in this organelle. Our results show that the specific inhibition of this system by injection of the monoclonal antibody causes a rapid intracellular increase of $[\text{Ca}^{2+}]_i$ and arrests further development. Calcium concentrations above 10^{-6}M cause disassembly of spindle microtubules both *in vitro*^{18,19} and *in vivo*^{5,6}. The dependence of the integrity of

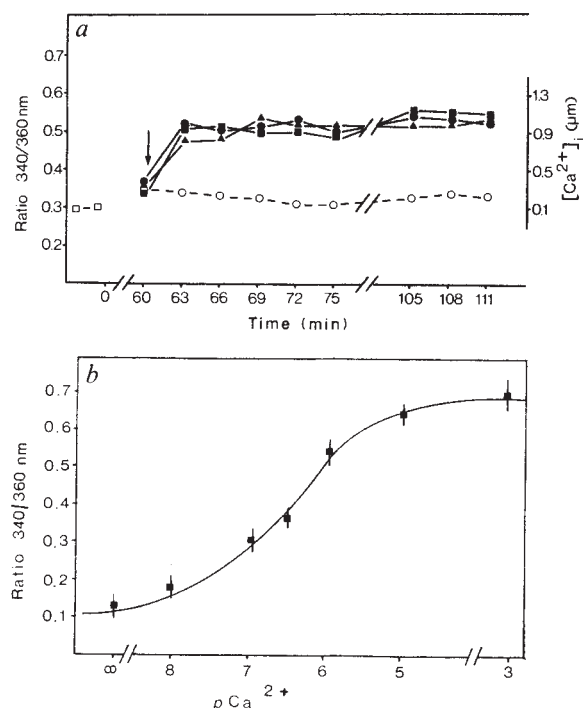


Fig. 3 Increase of $[Ca^{2+}]_i$ in a sea urchin blastomere after injection of antibody to 46K protein. *a*, Time course of the calcium increase. The injected cell was centred in the measuring field of the photometer. The data for the real Ca^{2+} on the right ordinate are derived from the calibration curve shown in *b*. Ca^{2+} is measured in the unfertilized egg (open squares), at metaphase shortly before microinjection (arrow), and after microinjection at the times indicated; the results of three different experiments are shown. $[Ca^{2+}]_i$ can be seen to rise slightly at mitosis, but increases sharply within two minutes after injection of the antibody solution. The course of the $[Ca^{2+}]_i$ fluctuations in a non-injected control cell is represented by the open circles. *b*, Relationship between the ratio of fura2-fluorescence (340/360 nm) and Ca^{2+} . *In vivo* calibration: sea urchin eggs were microinjected with fura2 to give final concentrations of 10–20 μ M. The surrounding sea water was replaced by a solution similar to the intracellular composition: 220 mM potassium glutamate, 500 mM glycine, 10 mM NaCl, 2.5 mM $MgCl_2$, pH 6.9, containing varying amounts of Ca^{2+} . The pCa^{2+} of this solution was controlled by varying the ratio of Ca-EGTA/EGTA (refs 20, 21). 1–2 μ M ionomycin were added to the injected eggs to allow equilibration of intra- and extracellular Ca^{2+} .

Methods. *Lytechinus pictus* eggs were microinjected with a solution of 15 mM fura2 in 120 mM potassium aspartate and 10 mM HEPES, pH 6.9, shortly before first mitosis²². Typical dye loadings were 10–20 μ M. Excitation of individual blastomeres was achieved by ultraviolet light from a mercury lamp (HBO 50) attached to a Zeiss microscope. Emitted light was measured by a photomultiplier. The excitation wavelength was alternated between 340 nm and 360 nm by using a wheel rotating at 1,500 r.p.m. and holding the respective interference filters, a bandpass on each of 2 nm (ref. 23). The characteristics of the dye are such that at an excitation wavelength of 340 nm there is an increase in emitted fluorescence intensity with increasing Ca^{2+} , and at the excitation wavelength 360 nm there is no change in the emitted fluorescence intensity (isosbestic point)²⁴. The ratios from the simultaneously recorded signals could then be interpreted using the equation²⁴

$$Ca^{2+} = K_d \frac{R_{340/360} - R_{min} \frac{S_{F2}}{S_{B2}}}{R_{max} - R_{340/360} \frac{S_{F2}}{S_{B2}}}$$

The K_d for fura2 was taken to be 774 nmol (ref. 9).

the mitotic apparatus on a functioning Ca^{2+} -transport system illustrated here suggests that the Ca^{2+} -sequestering membranes, which *in vivo* are sensitive to the antibody against 46K protein, are one of the key elements in the regulation of mitosis.

We thank Rainer Nobiling for access to his photometer and

for advice on calcium measurements and Signe Mähler for photography. Part of this work was supported by NIH to the Embryology Course, MBL, Woods Hole (M.H.) and by the Deutsche Forschungsgemeinschaft (C.P.).

Received 6 July; accepted 6 October 1987.

- Inoue, S. *J. Cell Biol.* **91** Suppl., 131–147 (1981).
- Hepler, P. K. & Wolniak, S. M. *Int. Rev. Cytol.* **90**, 169–236 (1984).
- Petzelt, C. & Hafner, M. in *Progress in Cell Biology* (eds, Schatten H. & Schatten G.) (Academic, in the press).
- Hepler, P. K. *J. Cell Biol.* **100**, 1363–1368 (1985).
- Kiehart, D. P. *J. Cell Biol.* **88**, 604–617 (1981).
- Izant, J. G. *Chromosoma* **88**, 1–10 (1983).
- Harris, P. *Expl. Cell Res.* **94**, 409–425 (1975).
- Silver, R. B., Cole, R. D. & Cande, W. Z. *Cell* **19**, 505–516 (1980).
- Poenie, M., Alderton, J., Tsien, R. Y. & Steinhardt, R. A. *Nature* **315**, 147–149 (1985).
- Keith, C. H., Maxfield, F. R. & Shelanski, M. L. *Proc. natn. Acad. Sci. U.S.A.* **82**, 800–804 (1985).
- Poenie, M., Alderton, J., Steinhardt, R. & Tsien, R. *Science* **233**, 886–889 (1986).
- Bennet, J. & Mazia, D. *Expl. Cell Res.* **134**, 494–498 (1981).
- Schantz, A. R. *J. Cell Biol.* **100**, 947–954 (1985).
- Petzelt, C. & Wüllfroth, P. *Cell Biol. int. Rep.* **8**, 823–840 (1984).
- Suprynowicz, F. & Mazia, D. *Proc. natn. Acad. Sci. U.S.A.* **82**, 2389–2393 (1985).
- Petzelt, C. & Hafner, M. *Proc. natn. Acad. Sci. U.S.A.* **83**, 1719–1722 (1986).
- Keith, C. H. *Cell Motility and the Cytoskeleton* **7**, 1–9 (1987).
- Weisenberg, R. C. *Science* **177**, 1104–1105 (1972).
- Salmon, E. D. & Segall, R. E. *J. Cell Biol.* **86**, 355–365 (1980).
- Swann, K. & Whitaker, M. *J. Cell Biol.* **103**, 2333–2342 (1986).
- Martell, A. E. & Smith, R. M. *Critical Stability Constants* Vol. 1, 269 (Plenum, London, 1974).
- Turner, P., Jaffe, L. A. & Fein, A. *J. Cell Biol.* **102**, 70–76 (1986).
- Nobiling, R. & Reiss, H. D. *Protoplasma* **139**, 20–24 (1987).
- Gryniewicz, G., Poenie, M. & Tsien, R. *J. biol. Chem.* **260**, 3440–3450 (1985).

HIV F/3' *orf* encodes a phosphorylated GTP-binding protein resembling an oncogene product

Bruno Guy*, Marie Paule Kieny*||, Yves Riviere†, Christian Le Peuch‡, Karin Dott*, Marc Girard§, Luc Montagnier† & Jean-Pierre Lecocq*

* Transgène S.A., 11 rue de Molsheim, 67082 Strasbourg Cedex, France

† Institut Pasteur, 25–28 rue du Dr. Roux, 75015 Paris Cedex, France

‡ U. 249 INSERM, Bvd. Henri IV, BP 5015, 34033 Montpellier Cedex, France

§ Pasteur Vaccins S.A., 1 Bd. R. Poincaré, 92430 Marnes-la-Coquette, France

Apart from the retroviral *gag*, *pol* and *env*^{1,2}, the HIV genome contains the *F* (3' *orf*) gene which encodes a polypeptide of 206 amino acids which is myristylated at the N-terminal and whose function is unknown^{3,4}. We have expressed the *F* gene in *Escherichia coli* and from a recombinant vaccinia virus, VVTGHIV. The F-protein produced in VVTGHIV-infected mammalian cells is myristylated, and is phosphorylated by protein kinase C at a residue close to the N-terminus like pp60-*src* (ref. 5). Purified bacterial F-protein also shows the GTPase, auto-phosphorylation and GTP-binding activities reported for the *ras* gene product^{6,7}. Furthermore, we show that expression of *F* in a CD4⁺ cell line down-regulates the CD4 (T4) antigen. These results suggest that *F* is important in the pathophysiology of AIDS (acquired immune deficiency syndrome).

The human immunodeficiency virus (HIV) genome contains a 3' open-reading frame *F* gene^{3,4,8} which is absent from the normal retroviral genome and is also found in HIV-2 and SIV^{9,10}. F-deficient HIV mutants show an altered cytopathic effect *in vitro*^{11,12}. The myristylation of F-protein³ at its N-terminal suggests it could be associated with the cell membrane, so we compared F with other membrane-associated proteins.

Sequence homologies were found between the N-termini of F-protein (ref. 8), the F homologue (*Rod*) of HIV-2 (ref. 9),

|| To whom correspondence should be addressed.

Fig. 1 *a*, Comparison of N-terminal sequences of proteins encoded by F-HIV-1_{Bru}, F-HIV-2_{Rod}, pp60-*src* and the membrane-proximal sequence of the *erb-B* encoded EGF receptor. Basic residues (arg/lys) are underlined. The arrow indicates the PKC phosphorylation site of the proteins from F-HIV-1_{Bru}, pp60-*src* and on the EGF receptor, and the putative PKC phosphorylation site¹⁴ on F-HIV-2-Rod. myr, Myristic acid; tm, transmembrane domain. The small arrow indicates the site on pp60-*src* phosphorylated by protein kinase A (ref. 5). The F-protein of HIV-2-Rod is myristylated (unpublished results). *b*, Comparison of the regions involved in GTP binding²¹ of v-Ha-*ras* (ref. 24), c-Ha-*ras* proteins (ref. 25) and α subunits of bovine G proteins *Gas* and *Gta2* (ref. 25), with the putative similar regions of F protein of HIV-1-Bru, HIV-2-Rod (ref. 9) and SIV (ref. 10). The arrow indicates the autophosphorylation site on thr-59 of v-Ha-*ras* (ref. 24). The region presented comprises part of the phosphoryl binding consensus site²³; additional localized homologies between *ras* and F are also present outside this region (not shown). Homologies between *ras* and F have previously been noted³⁵.

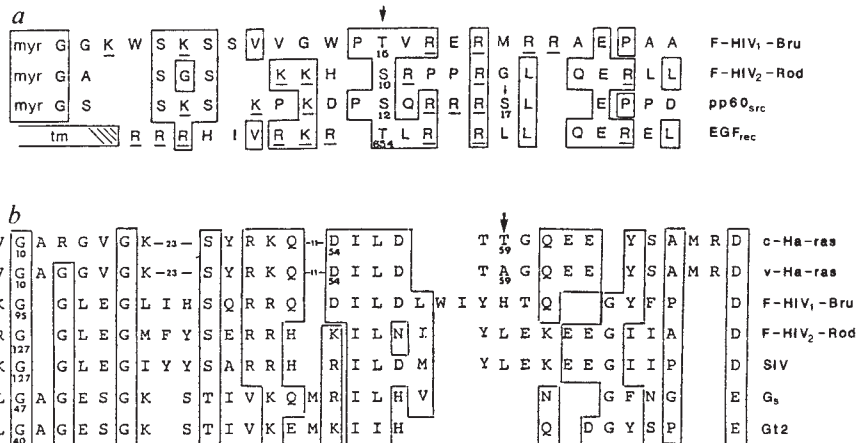
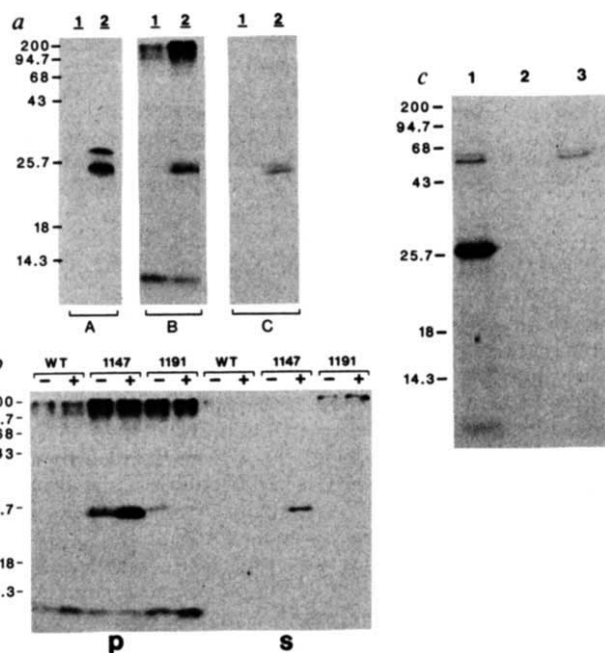


Fig. 2 Expression of F-protein. *a*, SDS-PAGE of radiolabelled proteins from vaccinia-infected hamster cells. Infecting viruses were (1) Copenhagen strain (2) VVTGfHIV-1147. *A*, [³⁵S]methionine labelling; *B*, ³²P_i labelling; *C*, ³H-myristic acid labelling. In *A* and *B*, extracts were immunoprecipitated using antiserum against F-protein, in *C*, cell extracts were analysed directly. *b*, Anti-F immunoprecipitates of BHK21 cells infected by vaccinia virus Copenhagen strain (WT), VVTGfHIV-1147 (1147) and VVTGfHIV-1191 (1191). Cells were labelled with ³²P_i and were either untreated (-) or treated with TPA (+). p, pellet; s, supernatant. *c*, Calcium- and phospholipid (PL)-dependent *in vitro* phosphorylation by PKC of F-protein produced in *E. coli*. Lane 1, F-protein + PKC + PL; lane 2, F protein + PL; lane 3, *E. coli* control + PKC + PL. *M_r* of standards (K) are shown to the left.

Methods. *a*, A *Bam*HI-*Hind*III fragment (nucleotides 8,032-9,173) of HIV ((LAV-1_{Bru} isolate) ref. 8) containing the *F* gene was cloned from plasmid PJ19-6 (ref. 20) into a M13 vector opened at *Kpn*I and *Hind*III sites, using a *Bam*HI-*Kpn*I adaptator 5' GATCGTAC 3' (M13TG170). A *Bam*HI site was created upstream from the initiation codon of the *F* gene by localized mutagenesis using the oligonucleotide 5' GAAAGGATCCTGCTATAAG 3' and the *Bam*HI-*Sst*I fragment was cloned into plasmid pTG186-poly for the transfer into the vaccinia virus genome³⁶ to generate VVTGfHIV-1147. F-protein was immunoprecipitated using serum from an HIV-seropositive individual, after labelling of infected BHK-21 cells for 4 h. [³⁵S]methionine-labelled and orthophosphate-labelled extracts were immunoprecipitated^{36,37}. For myristate-labelling an aliquot of the cell lysate was loaded onto the SDS gel without immunoprecipitation. *b*, Oligonucleotide 5' TCTTCCCTCACTGCAGGCCATCC 3' was used to mutate the sequence coding for Threonine 15 of the F-protein to Alanine in M13TG173 (see Fig. 1), generating M13TG1151. After cloning of the *Bam*HI-*Sst*I fragment into pTG186-poly, and construction of the recombinant VVTGfHIV-1191, 10⁶ BHK-21 cells were infected with 0.1 PFU/cell for 12 h and labelled for 4 h with 100 μ Ci of carrier free ³²P_i. TPA (Sigma) (100 ng per 10⁶ cells) was added 10 min before collecting the cells and immunoprecipitation³⁷ with HIV seropositive serum. *c*, Oligonucleotide 5' GCCACCCATAG-GATCCCCAAATCCTTT 3' was used to generate a *Bam*HI site immediately upstream of the ATG initiation codon of the *F* gene in recombinant M13TG170, to give M13TG1106. The *Bam*HI-*Hind*III fragment containing *F* was cloned in *E. coli* plasmid pTG959 under the control of the thermoinducible P_L promoter of λ to generate pTG1166. Induction at 42 °C leads to the expression of F-protein to ~20% of total cell protein (not shown). F-protein was purified to 75% from the insoluble fraction by solubilization in 0.2% SDS; SDS was removed by addition of KCl (final concentration 75 mM) and centrifugation. The supernatant was dialysed extensively against 10 mM NaH₂PO₄, pH 7.4. PKC was purified from rat brain to homogeneity²⁰. 30 μ l of purified PKC was added to 20 μ l solution containing 5 mg ml⁻¹ of purified F-protein, 80 μ g ml⁻¹ of phosphatidyl serine, 8 μ g ml⁻¹ dialeline, 20 mM HEPES-NaOH, pH 7.4, 10 mM MgCl₂, 0.5 mM CaCl₂, 7 mM DTT, 0.7 mM ATP, [³²P]ATP (2,000 c.p.m. pmole⁻¹). Incubation was performed at 25 °C for 30 min and the reaction stopped by addition of 50 μ l 5 M urea, 0.2 M SDS, 0.5 M dithiothreitol, pH 6.8. Proteins were separated by SDS-PAGE (15% acrylamide), dried and exposed for autoradiography. The kinetics of incorporation show that at least 0.2 moles of [³²P]ATP per mole of F-protein are incorporated in 30 min in a Ca²⁺ phospholipid-dependent reaction (not shown).

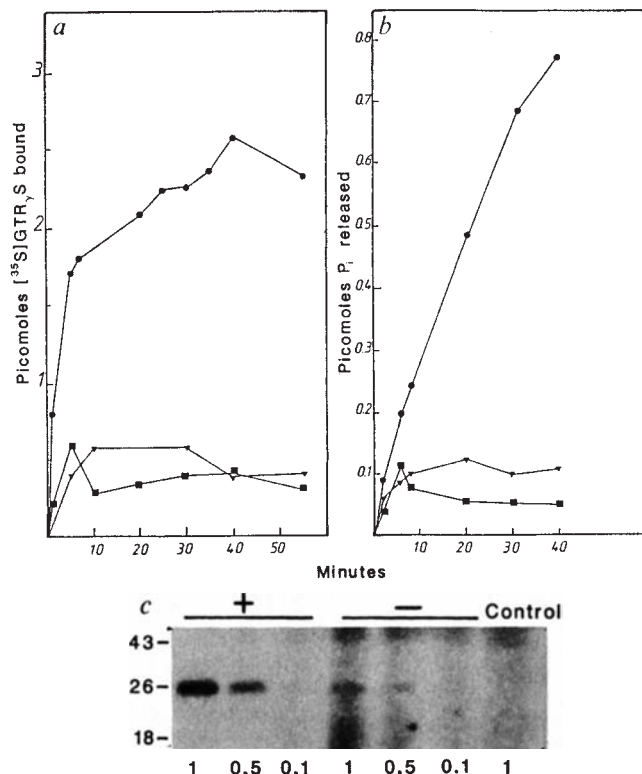


pp60-*src* protein⁵ and the epidermal growth factor (EGF) receptor encoded by the *erb-B* oncogene¹³ (Fig. 1a), all of which contain a common site for phosphorylation by protein kinase C (PKC) (ref. 14). To determine whether F-protein is phosphorylated *in vivo* at this sequence we constructed a recombinant vaccinia virus, VVTGfHIV-1147, which expresses the F-protein of a cloned HIV provirus (LAV-Bru) (ref. 8, Fig. 2 legend). Radiolabelled proteins from VVTGfHIV-infected cells were tested for reaction with AIDS sera: ~30% of human HIV-positive sera specifically immunoprecipitated two protein

species from infected cell extracts. The apparent relative molecular mass (*M_r*) of these proteins was 25,000 (25 K) and 27 K (p25-F and p27-F) (Fig. 2a), as previously reported¹⁵ for F-protein from HIV-infected cells. Only the p25-F species was labelled when ³H-myristic acid was added to the cell culture medium (Fig. 2a); p27-F is presumed to be the non-myristylated F polypeptide.

To assess phosphorylation of the recombinant F-protein VVTGfHIV-1147-infected cells were incubated with ³²P_i and the labelled proteins tested for reaction with AIDS sera. Two closely

Fig. 3 *a*, GTP binding assay. Rates of [γ - 35 S] thioGTP binding to F protein in the presence (●) or absence (■) of Mg^{2+} ions, and in a solution containing an unrelated protein in the presence of Mg^{2+} ions (▼). *b*, GTPase assay. Amounts of $^{32}P_i$ released in samples containing F-protein in the presence (●) or absence (■) of Mg^{2+} ions, and samples containing an unrelated protein (▼) in the presence of Mg^{2+} ions. *c*, Autokinase assay. Migration of phosphorylated protein on SDS-PAGE after 60 min incubation with [γ - ^{32}P]GTP at 37 °C. Reactions were with F protein in the presence (+) or absence (–) of Mg^{2+} ions, control protein in presence of Mg^{2+} ions. Loading of proteins was 2 μ l of amounts indicated in μ g in 20 μ l. M_r of standards are shown to the left (K). **Methods.** *a*, GTP binding assays (40 μ l) were carried out as described³⁸ in the presence or absence of Mg^{2+} ions, with 15 picomoles of [γ - 35 S]-thioGTP (New England Nuclear) (10^5 c.p.m. μ mol $^{-1}$) and 0.5 μ g purified F protein. The amount of F in each assay was estimated as 20 picomoles using apparent M_r 25K. At the times indicated samples were filtered on nitrocellulose (HA 0.45 μ M Millipore), rinsed and the retained radioactivity determined by liquid scintillation counting. A control protein over-produced in *E. coli* (M_r 28K) and purified in parallel showed no activity in this assay. *b*, GTPase activity was determined⁶ using 2.5 mM unlabelled ATP to suppress non-specific ATP/GTPase activity, 1 μ M [γ - ^{32}P] GTP (10^5 c.p.m. μ mol $^{-1}$), and 0.5 μ g protein (estimated to 20 pmol). *c*, Autokinase activity was determined⁶ with the same proteins in 20 μ l buffer containing 10 μ M [α - ^{32}P] GTP (10^5 c.p.m. μ mol $^{-1}$). Protein staining shows that the radio-labelled band comigrates with F (not shown). Experiments performed with [α - ^{32}P] GTP or [γ - ^{32}P] ATP were negative. Radioactivity in labelled F protein was 1,800 c.p.m. for 200 ng loaded on the gel.



migrating 25K phosphoproteins, pp25-F₁ and pp25-F₂, were identified by this procedure (Fig. 2a); addition of excess unlabelled purified F-protein produced from *E. coli* (below) to the immunoprecipitation reaction eliminated the labelled bands (not shown). To investigate whether F-phosphorylation occurs at the protein kinase C consensus site¹⁴, *thr*-15 (Fig. 1a), we constructed a second recombinant vaccinia virus, VVTGfHIV-1191, in which the codon for *thr*-15 was replaced by a codon for alanine (Fig. 2 legend), a substitution occurring naturally in certain HIV isolates^{16,17}. Infection of cultured cells with the *ala* recombinant resulted in synthesis of both the 25K and 27K species; the 25K species was specifically labelled with 3H -myristic acid, showing that the *thr*-*ala* mutation does not affect N-terminal myristylation (not shown). Extracts of $^{32}P_i$ -labelled cells infected with VVTGfHIV-1191 were then tested for reaction with AIDS sera. The *thr*-*ala* substitution resulted in loss of pp25-F₁ while pp25-F₂ remained unaffected, demonstrating that pp25-F₁ is predominantly phosphorylated at *thr*-15. Alkaline lability of phosphorylated pp25-F₂ suggests that phosphorylation is at an unidentified serine residue¹⁸ (not shown).

To investigate whether F-protein phosphorylation is mediated by PKC we labelled infected cells with $^{32}P_i$ in the presence of the protein kinase C activator⁵, 12-O-tetradecanoyl-phorbol-13-acetate (TPA) (ref. 19). Phosphorylation of pp25-F₁ was markedly stimulated by TPA (Fig. 2b) and labelling was eliminated by the *thr*-*ala* substitution. pp25-F₂ phosphorylation was unaffected by TPA. We then expressed the F coding sequence in *E. coli* and purified the product to >75% homogeneity (legend Fig. 2). Purified bacterial F-protein was found to be efficiently phosphorylated *in vitro* by PKC in a Ca^{2+} - and phospholipid-dependent reaction²⁰, confirming that F is a specific substrate for protein kinase C (Fig. 2c). F-protein activity of some HIV isolates could thus be regulated by PKC phosphorylation like other target proteins^{13,21}.

Similarities between F and cellular oncogene products possessing kinase activity^{5,13}, led us to examine the F amino-acid sequence for a potential nucleotide-binding site. This analysis revealed a homology with regions involved in the GTP-binding site^{22,23} of the *ras* family of proteins^{24,25} and other GTP-binding proteins²⁵ (Fig. 1b). To explore whether F could

bind GTP like *ras*, bacterial F was incubated with [γ - 35 S]thioGTP and assessed for nucleotide binding by a filter-retention assay; F-protein was found to bind GTP specifically (Fig. 3a) and we determined whether F might exhibit similar properties. GTP binding of v-Ha-*ras* is associated with autophosphorylation and GTPase activity⁶. Purified F protein was incubated with [γ - ^{32}P] GTP and analysed by gel electrophoresis and autoradiography (Fig. 3c): bacterial F-protein was found to be phosphorylated under these conditions. A serine residue is implicated in the phosphorylation because of the lability in alkali¹⁸ of the phosphoprotein (not shown). Phosphorylation of this residue could generate the pp25-F₂ species observed *in vivo* (see above).

To determine whether F also possesses GTPase activity, we measured the rate of GTP cleavage by purified bacterial F-protein, when it was found that GTP was cleaved at about the same rate as reported⁶ for purified v-Ha-*ras* protein (1 mMol GTP per mol protein per minute for F; 0.2 mMol GTP per mol protein per minute for *ras*). The kinetics of the autophosphorylation reaction were also compared, when F was found to incorporate 0.17 mMol P_i per mol F-protein per minute, in good agreement with the value⁶ for v-Ha-*ras* (0.33 mMol P_i per mol protein per minute).

These common features of F-protein and the v-Ha-*ras* and pp60-*src* proteins suggest that F could intervene in cell regulation. There is to date no evidence that HIV-infected cells are subject to oncogenic transformation, but activation of cellular regulatory pathways, for instance by phorbol ester treatment, has been reported to down-regulate the lymphocyte T4 antigen²⁶, an important biological event in HIV infection^{27–29}. We therefore investigated whether F-protein could be involved in the down-regulation of T4 observed in HIV-infected cells. Fluorescence-activated cell sorting showed that the T4 antigens (Ieu3a and OKT4 markers) on CEM T4 cells³⁰ infected with either the *Thr* or *Ala* variants of VVTGfLAV were markedly reduced; this reduction was not observed in uninfected cells or in cells infected with a control recombinant (Fig. 4).

We have shown that the F-protein of HIV is a myristylated GTP-binding phosphoprotein with features similar to the cellular *src* and *ras* oncogene products. It is interesting that F

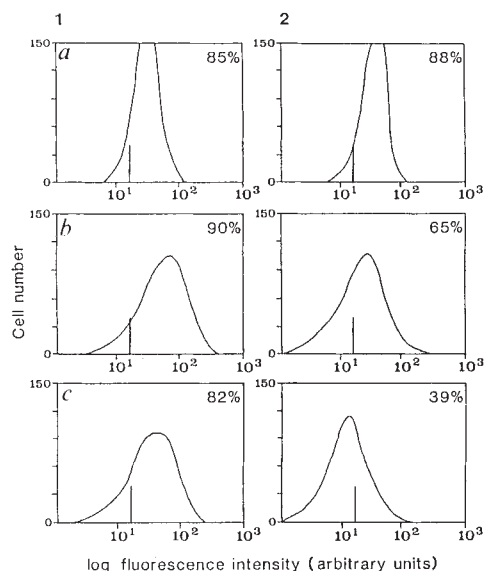


Fig. 4 Flow cytometry of vaccinia-infected CEM-T4 (ref. 30) cells labelled with antibodies to lymphocyte antigens. Infecting viruses were a control recombinant expressing an unrelated protein coding sequence (1) and VVTG/HIV-1191 (2); *a*, infected cells were incubated with monoclonal antibodies 4B4, *b*, leu3a and *c*, OKT4. The % fluorescent cells is indicated for each experiment. Reactions carried out with mouse antivaccinia polyclonal antibodies indicated that all the cells were infected by the recombinant viruses (not shown). Note that antigen down-regulation appears to depend on the amount of F-protein expression; no effect on T4 expression was observed 2 h after infection whereas the down-regulation is most pronounced at 16 h (not shown). Infection by a recombinant vaccinia virus expressing the *env* gene also leads to significant down-regulation of the T4 leu3a marker but the T4-OKT4 marker is less affected (Y. R. *et al.*, manuscript in preparation); down-regulation by *env* could be due to epitope masking rather than to disappearance of the T4 molecule from the cell surface.

Methods. 3×10^7 CEM-T4 cells³⁰ were infected with recombinant viruses (50 PFU cell^{-1}) in 1 ml PBS. After 1 h at room temperature, cells were diluted in 10 ml RPMI medium (Gibco) and 16 h later 10^6 cells ($100 \mu\text{l}$) were incubated for 30 min at 4°C with: murine monoclonal antibodies against CD4 (1/100), OKT4 (Ortho) and Leu3a (Becton Dickinson), or the 4B4 murine monoclonal antibody (Coultronics) directed against a 135K cell-surface protein unrelated to CD4. Cells were washed twice and resuspended in $100 \mu\text{l}$ of a biotinylated sheep anti-mouse immunoglobulin or anti-human immunoglobulin (Amersham), 1/100 for 30 min at 4°C . Cells were washed and resuspended into $20 \mu\text{l}$ of streptavidin-phycoerythrin (Becton Dickinson) (1/50) for 30 min at 4°C . After washing, cells were resuspended in 0.4% formaldehyde in PBS and analysed with a fluorescence-activated cell sorter (Becton Dickinson). Incubations, washes and dilutions of sera were in Ca^{2+} and Mg^{2+} -free PBS supplemented with 1% bovine serum albumin and 0.1% NaN_3 .

resembles both proteins, for *ras* may mediate signal transduction by *src*³¹. We have shown that F may participate with *env* (refs 28, 29) in down-regulation of T4 lymphocytes. It is of note that, in contrast to *gag* and *env*, F synthesis continues in the absence of *art* protein³², and could be responsible for the diminished T4 expression observed^{33,34} at the surface of latently infected cells that do not express HIV structural proteins. Latent infection of lymphocytes is important to the pathophysiology of AIDS, and we speculate that F-protein, by intervening in cellular regulatory pathways, may play a key role in the establishment or maintenance of latent HIV infection.

We thank R. Lathe for his contribution to this work and R. Olivier for help with flowcytometry; also A. Laurent and M. Basset for help with *in vitro* PCK phosphorylation, F. Jaeger and S. Chamaret for technical assistance, and F. Daul, E. Chambon and N. Monfrini for preparation of the manuscript. We thank A. Capron, R. Drillien, D. Spehner and D. Stehelin for

discussions, and E. Eisenmann, P. Chambon and P. Kourilsky for their continued interest. This study was supported by Pasteur Vaccins.

Received 4 June; accepted 7 October 1987.

1. Wong-Staal, F., Pranab, K. L. & Ghayeb, J. *AIDS Res. hum. Retrovir.* **3**, 33–39 (1987).
2. Mitsuya, H. & Broder, S. *Nature* **325**, 773–778 (1987).
3. Allan, S. J. *et al. Science* **230**, 810–813 (1985).
4. Arya, S. K. & Gallo, R. C. *Proc. natn. Acad. Sci. U.S.A.* **83**, 2209–2213 (1986).
5. Gould, K. L. *et al. Cell* **42**, 849–857 (1985).
6. Hattori, S., Uish, L. S., Halliday, K. & Shihe, T. Y. *Molec. cell. Biol.* **5**, 1449–1455 (1985).
7. Trahey, M. *et al. Molec. cell. Biol.* **7**, 541–544 (1987).
8. Wain Hobson, S., Sonigo, P., Danos, O., Cole, S. & Alizon, M. *Cell* **40**, 9–17 (1985).
9. Guyader, M. *et al. Nature* **326**, 662–669 (1987).
10. Chakrabarti, L. *et al. Nature* **328**, 543–547 (1987).
11. Terwilliger, E., Sodroski, J. G., Rosen, C. A. & Haseltine, W. J. *J. Virol.* **60**, 754–760 (1986).
12. Luciw, P. A., Cheng-Mayer, C. & Levy, J. A. *Proc. natn. Acad. Sci. U.S.A.* **84**, 1434–1438 (1987).
13. Hunter, T., Ling, N. & Cooper, J. A. *Nature* **311**, 480–483 (1984).
14. Woodgett, J. R., Gould, K. L. & Hunter, T. *Eur. J. Biochem.* **161**, 177–184 (1986).
15. Franchini, G., Robert-Guroff, M., Ghayeb, J., Chang, N. T. & Wong-Staal, F. *Virology* **155**, 593–599 (1986).
16. Ratner, L. *et al. Nucleic Acids Res.* **13**, 8219–8229 (1985).
17. Alizon, M., Wain-Hobson, S., Montagnier, L. & Sonigo, P. *Cell* **46**, 63–74 (1986).
18. Cooper, J. A., Sefton, B. M. & Hunter, T. *Meth. Enzym.* **99**, 387–402 (1983).
19. Nishizuka, Y. *Science* **233**, 305–312 (1986).
20. Le Peuch, C. J., Ballester, R. & Rosen, O. M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 6858–6862 (1983).
21. Bollag, G. E., Roth, R. A., Beaudoin, J., Nochly-Rosen, D. & Koshland, D. E. *Proc. natn. Acad. Sci. U.S.A.* **83**, 5822–5824 (1986).
22. Sigal, I. S., Gibbs, J. B., D'Alonzo, J. S. & Scolnick, E. *Proc. natn. Acad. Sci. U.S.A.* **83**, 4725–4729 (1986).
23. Dever, T. E., Glynnias, M. J. & Marwick, W. C. *Proc. natn. Acad. Sci. U.S.A.* **84**, 1814–1818 (1987).
24. Dhar, R. *et al. Science* **217**, 934–937 (1982).
25. Gilman, A. G. *et al. Rev. Biochem.* **56**, 615–649 (1987).
26. Dröge, W. *Immun. Today* **7**, 340–343 (1986).
27. Seligmann, M. *et al. New Engl. J. Med.* **311**, 1286–1292 (1984).
28. Hoxie, J. A. *et al. Science* **234**, 1123–1127 (1986).
29. Klatzmann, D. *et al. Nature* **312**, 767 (1984).
30. Gluckman, J. C., Klatzmann, D. & Montagnier, L. *et al. Rev. Immun.* **4**, 97–117 (1986).
31. Smith, M. R., De Gubicus, S. & Stacey, D. W. *Nature* **320**, 540–543 (1986).
32. Feinberg, M. B., Janet, R. F., Aldovini, A., Gallo, R. C. & Wong-Staal, F. *Cell* **46**, 807–817 (1986).
33. Folks, T. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 4539–4543 (1985).
34. Hoxie, J. A., Haggarty, B. S., Rackowski, J. L., Pillsbury, N. & Levy, J. A. *Science* **229**, 1400–1402 (1985).
35. Samuel, K. P., Seth, A., Konopka, A., Lautenberger, J. A. & Papas, T. S. *FEBS Lett.* **218**, 81–86 (1987).
36. Kieny, M. P. *et al. Nature* **312**, 163–166 (1984).
37. Acres, R. B., Carlson, P. J., Mochizuki, D. Y. & Gallis, B. *J. biol. Chem.* **261**, 16210–16214 (1986).
38. Sternweis, P. C. & Robishaw, J. D. *J. biol. Chem.* **259**, 13806–13813 (1984).

Novel source of 1,2-diacylglycerol elevated in cells transformed by Ha-ras oncogene

Juan Carlos Lacal, Jorge Moscat* & Stuart A. Aaronson

Laboratory of Cellular and Molecular Biology, National Cancer Institute, Building 37, Room 1E24, Bethesda, Maryland 20892, USA

Genes involved in the transduction of signals required for normal cell proliferation commonly appear to be subverted in the neoplastic process¹. One such group is the highly conserved family of *ras* genes, which have been detected as transforming genes in a wide variety of naturally occurring tumours. By analogy with other known G proteins², the p21 proteins encoded by *ras* genes may act as regulatory proteins in the transduction of signals that lead to DNA synthesis. A major pathway involved in the DNA synthesis induced by growth factors is mediated by phosphatidylinositol turnover: cleavage of phosphoinositides by phospholipase C produces 1,2-diacylglycerol, and inositol phosphates³. The former acts as an essential cofactor for protein kinase C (ref. 4), and inositol-(1,4,5)-triphosphate mobilizes Ca^{2+} from non-mitochondrial intracellular stores⁵. We demonstrate a reproducible increase in 1,2-diacylglycerol, in the absence of a detectable increase in inositol phosphates, in transformed cells containing Ha-ras

* Permanent address: Departamento Medicina Experimental, Hospital Provincial de Madrid, Spain.

oncogenes and with different membrane targeting signals for the *ras* p21 protein. These findings suggest that a source other than phosphoinositides exists for the generation of 1,2-diacylglycerol and that the *Ha-ras* oncogene specifically activates this novel pathway for 1,2-diacylglycerol production.

Most *ras* oncogenes are activated by point mutations at one of two main sites in their coding sequences⁶. The p21 proteins encoded by *ras* genes have a relative molecular mass (M_r) of 21,000 (21K), bind guanine nucleotides⁷, have intrinsic GTPase activity^{8,9}, and function by attachment to the inner cell membrane¹⁰. We have generated a series of *Ha-ras* oncogene constructs lacking the carboxy terminal sequences which are responsible for the attachment of the p21 protein to the membrane by means of post-translational acylation with fatty acids. These constructs fail to transform NIH/3T3 cells (J.C.L. *et al.*, manuscript in preparation). In contrast, chimeric *Ha-ras* genes are fully transforming when they possess sequences encoding the N-terminal 15 amino acids of the *v-src* gene product which is responsible for myristylation and targeting pp60^{v-src} to the inner surface of the cell membrane. We have investigated the effects of functional *ras* oncogene products with different membrane targeting signals on phospholipid metabolism using these mutants.

The pLTR-41 vector encodes the wild-type transforming *Ha-ras* p21 protein, and pLTR-p18d encodes a non-transforming protein with amino acids 164–189 were deleted and replaced by a stretch of six unrelated amino acids. Vectors pSR-41 and pSR-p18d contain coding sequences for the 15 terminal amino acids of pp60^{v-src}, fused inframe to the initiation codon for the p21 proteins encoded by pLTR-41 and pLTR-p18d, respectively. Each of these latter plasmids can transform efficiently, indicating that p21 oncogene activity is independent of its 25 carboxy terminal amino acid residues, provided that the protein is translocated to the membrane by either method of fatty acid acylation. Moreover, the morphology of the phenotype of these transformants was indistinguishable from that of the wild-type *ras* transformants (J.C.L. *et al.*, manuscript in preparation).

We initially investigated the concentrations of inositol phosphates (Ins Ps) and 1,2-diacylglycerol (DAG) in control NIH/3T3 cells and in NIH/3T3 lines containing each of the vectors. As seen in Table 1, quiescent NIH/3T3 cells stimulated with serum for one hour showed a significant increase in both DAG and Ins Ps, consistent with phosphatidyl inositol 4,5-bisphosphate (PtdIns P2)-specific phospholipase C (PLC) activation. Increase in Ins Ps was also observed following PDGF (Platelet-derived growth-factor) stimulation, in agreement with previous results^{5,11,12}. Wild-type *Ha-ras* transformants containing pLTR-41 also gave a substantial elevation in DAG compared with control NIH/3T3 cells, so did transformants containing pSR-41 or pSR-p18d (Table 1). In contrast, NIH/3T3 cells containing the non-transforming pLTR-p18d plasmid or pSV₂ *neo* showed a negligible variation in DAG concentration.

Although transformants induced by *ras* oncogene products with different membrane targeting mechanisms both showed an increase in DAG, there was no detectable elevation of Ins Ps such as would be expected should the primary DAG source be the hydrolysis of PtdIns Ps by PLC. The pSR-41 and pSR-p18d transformants, for example, caused Ins Ps to actually decrease relative to amounts in quiescent or serum- or PDGF-stimulated NIH/3T3 cells (Table 1). These results were reproducible when independent cultures containing the same *ras* gene construct were analysed (Table 1). NIH/3T3 cells transformed by the *v-sis* oncogene encoding a PDGF homologue¹³ (*v-sis*/PDGF-2 transformant), however, gave a fourfold increase in Ins Ps, in agreement with results from PDGF-stimulated NIH/3T3 cells.

Phospholipids other than PtdIns Ps could be sources of DAG under physiological conditions^{14–18}. We therefore investigated the steady state amounts of phosphorylcholine, phosphorylethanolamine and phosphorylserine resulting from phosphodiester attack by PLC phosphatidyl choline (Ptd Cho), phos-

Table 1 Steady-state levels of [2-³H]glycerol-labelled DAG and of myo-[2-³H] inositol-labelled Ins Ps in different *H-ras* transformants

NIH/3T3 cell line	Transformed phenotype	PLC products	
		Ins Ps	DAG
Control	–	1.00	1.00
+serum (1 h)	–	1.40 ± 0.05	1.25 ± 0.08
+PDGF (1 h)	–	3.00 ± 0.02	ND
Transfectant containing:			
pLTR-41 no. 1	+	1.00 ± 0.01	1.53 ± 0.13
no. 2	+	1.02 ± 0.03	1.45 ± 0.14
pSR-41	+	0.61 ± 0.03	1.45 ± 0.05
pSV ₂ - <i>neo</i>	–	0.90 ± 0.06	1.08 ± 0.06
pLTR-p18d + pSV ₂ - <i>neo</i>	–	0.81 ± 0.03	1.13 ± 0.03
pSR-p18d	+	0.38 ± 0.03	1.38 ± 0.03
<i>v-sis</i> /PDGF-2	+	4.2 ± 0.09	ND

Plasmids included pLTR-41, which carries the wild-type *v-H-ras* oncogene, and pLTR-p18d, which deletes the codons 165–189 of *v-H-ras* and replaces them with a stretch of six unrelated codons. pSR-41 and pSR-p18d contain *v-H-ras* or p18d fused in-frame at their respective initiation codons with a synthetic linker encoding the N-terminal 15 amino acids of the avian p60^{v-src} protein (J.C.L. *et al.*, manuscript in preparation). Cell lines containing each transforming construct as well as pSV₂-*neo* cotransfected cell lines for the non-transforming pLTR-18d construct were used. Cells were grown in Dulbecco's modified Eagle's medium supplemented with 10% calf serum. Cultures were labelled to equilibrium for 70 h in serum-containing medium in 35 mm dishes either with 10 μ Ci of [2-³H]glycerol (1 Ci mol^{–1}; Amersham) or myo-[2-³H]inositol (10–20 Ci mmol^{–1}; Amersham) per dish. After radiolabelling, cultures were washed and maintained for 4 h in serum-free medium, after which the medium was removed and cells incubated for 1 h in 1 ml of serum-free medium or serum-containing medium. In experiments designed for determination of Ins Ps release, LiCl (20 mM) was present during the 1 h incubation period. For DAG determination, lipids were extracted and fractionated by thin layer chromatography (TLC) as described³⁰. [2-³H]glycerol-labelled DAG production was normalized for the amount of radioactivity remaining at the origin of the TLC plate, which corresponds to total phospholipids. For Ins Ps analysis, reactions were terminated and total Ins Ps were eluted from Dowex 1 × 8 columns with 1.2 M ammonium formate/0.1 M formic acid as described previously³¹. These data were normalized for the radioactivity content in the TCA-insoluble material. The recovery was at least 80% of the precursor used in each experiment. Results are expressed as stimulation over control values for NIH/3T3 cells, and represent the mean values of three independent experiments with incubation in duplicate. ND, not determined.

phatidyl ethanolamine (Ptd Etn), and phosphatidyl serine (Ptd Ser), respectively. As shown in Table 2, an increase of 40% to 147% in the release of phosphorylcholine and an increase of 60% to 130% in the release of phosphorylethanolamine was observed in independent transformants containing the wild-type viral *Ha-ras* gene (pLTR-41), as well as in the cell lines containing transforming chimeric *ras* genes, pSR-41 or pSR-p18d. As controls, cells containing the non-transforming pLTR-p18d construct had phosphorylcholine and phosphorylethanolamine in concentrations similar to NIH/3T3 cells (Table 2). Moreover, NIH/3T3 cells stimulated by addition of 10% calf serum showed no increase in the release of these metabolites (Table 2), although the production of Ins Ps was stimulated under similar conditions (Table 1). Finally, no significant differences in amounts of phosphorylserine were observed in control, serum-stimulated, or wild-type *Ha-ras* oncogene-transformed NIH/3T3 cells (Table 2).

The specificity of the phosphorylcholine and phosphorylethanolamine increase associated with the *Ha-ras* transformed state is illustrated by *v-sis*/PDGF-2 transformant cells cultured under identical conditions: in contrast to the results with *Ha-ras* oncogene transformants, concentrations of both phosphorylcholine and phosphorylethanolamine were substantially reduced (Table 2). These results and the different observations on Ins Ps generation in the two systems indicate that the growth pathways

Table 2 Steady state levels of phosphorylcholine (P-Cho), phosphorylethanolamine (P-Eth) and phosphorylserine (P-Ser) in different *ras* transformants

NIH/3T3 cell line	Transformed phenotype	P-Cho	PLC products P-Eth	P-Ser
Control	—	1.00	1.00	1.0
+serum (1 h)	—	1.00 ± 0.01	1.00 ± 0.01	1.0 ± 0.02
+PDGF (1 h)	—	0.77 ± 0.13	ND	ND
Transfectant containing:				
pLTR-41 no. 1	+	2.47 ± 0.02	2.30 ± 0.1	1.0 ± 0.01
no. 2	+	2.08 ± 0.09	1.70 ± 0.08	1.0 ± 0.03
pSR-41	+	1.62 ± 0.05	1.90 ± 0.02	ND
pSV ₂ -neo	—	0.87 ± 0.03	0.83 ± 0.05	ND
pLTR-p18d + pSV ₂ -neo	—	0.81 ± 0.02	0.91 ± 0.07	ND
pSR-p18d	+	1.42 ± 0.02	1.60 ± 0.05	ND
v-sis/PDGF-2	+	0.40 ± 0.04	0.40 ± 0.03	ND

NIH/3T3 cells were transfected with the constructs as described in legend to Table 1. The molecularly cloned simian sarcoma virus genome³² was the source of v-sis/PDGF-2 (ref. 33). Cultures were labelled to equilibrium in serum-containing medium for 70 h in 33 mm dishes with 10 μ Ci per dish of either [methyl-³H] choline (70–90 Ci ml⁻¹; Amersham) or [2-¹⁴C]-ethanol-1-ol-2-amine (40–60 mCi ml⁻¹; Amersham) or L-[U-¹⁴C] serine (>150 mCi ml⁻¹; Amersham). The experimental protocol followed was identical to that described in Table 1. Reactions were terminated by TCA precipitation as described³¹. Water-soluble PLC products were fractionated by TLC by using silica gel G and the following solvent system: methanol/0.5% NaCl/NH₃ (100/100/2; v/v/v). Samples were run with the corresponding standards. These were obtained by deacylation of the corresponding commercially available labelled phospholipids³⁴. Phosphorylases were obtained by glycerol removal of the deacylated compounds³⁴. The results are expressed as described in the legend to Table 1. ND, not determined.

activated by *sis* and *ras* oncogenes differ in their effects on phospholipid metabolism.

Investigation of these pathways has intensified following evidence of the frequent involvement of *ras* oncogenes in naturally occurring malignancies. Microinjection of mammalian transforming *ras* p21 proteins expressed in bacteria into NIH/3T3 cells leads to a potent stimulation of DNA synthesis^{8,19} and induces a rapid rise in intracellular pH²⁰. This pH change is mediated by the amiloride-sensitive Na⁺/H⁺ antiporter, which is activated by PKC (ref. 21), and therefore supports the conclusion that PKC is activated early in response to Ha-*ras* p21. Finally, phorbol esters, which down-regulate PKC (refs 22, 23), inhibit the mitogenic response to microinjected *ras* p21, an inhibition that can be overcome by PKC microinjection²³. All these findings argue that the mitogenicity of the *ras* oncogene product is mediated significantly by PKC activation.

The increased DAG levels observed in *ras*-transformed cells were independent of the mechanism used to translocate *ras* p21 to the membrane. Microinjection of *ras* p21 may stimulate phospholipase A₂ activity²⁴. Other reports have indicated increases in Ins Ps as well as in DAG in *ras* transformants, causing speculation that *ras* p21 could act as the G regulatory protein controlling PtdIns Ps-specific PLC^{12,25–28}. But our studies and others²⁶, detect no increase in Ins Ps in *ras* transformants. Instead, we found elevation of both phosphorylcholine and phosphorylethanolamine which arise with DAG from PLC attack on two major phospholipid constituents of the cell membrane, Ptd Cho and Ptd Etn. Thus, we propose that Ptd Cho and Ptd Etn, rather than PtdIns Ps, are the major sources of elevated DAG in *ras*-transformed cells.

Relatively little is known concerning the physiological importance of the phosphodiester hydrolysis of Ptd Cho or Ptd Etn in mammalian cells. After agonist challenge to hepatocytes DAGs with a different fatty acid component from PtdIns Ps were generated¹⁶, suggesting that DAG sources other than PtdIns Ps could exist under physiological conditions. Bombesin stimulation of fibroblasts is also associated with increased amounts of phosphorylcholine, which implies that this growth factor affects Ptd Cho phosphodiester hydrolysis¹⁷. G proteins have been implicated in Ptd Cho hydrolysis as a result of increased phosphorylcholine release observed in isolated hepatocyte membranes in response to an agonist, an effect enhanced by non-hydrolysable GTP analogues¹⁸. Whether Ha-*ras* p21 directly or indirectly activates a PLC specific for Ptd

Cho and Ptd Etn through its function as a G protein remains to be determined.

In serum- or PDGF-stimulated NIH/3T3 cells or in cells chronically expressing the v-sis/PDGF-2 oncogene, the chief alterations in phospholipid metabolism involved increased phosphoinositide turnover. A rapid but transient rise in phosphorylcholine release has been reported to follow serum stimulation of a pre-adipocyte cell line²⁹; we did not detect any increase in this metabolite in serum- or PDGF-stimulated NIH/3T3 cells even at early time points (data not shown). In fact, PDGF-treated and v-sis/PDGF-2 transformed cells both had decreased amounts of the Ptd Cho metabolites. Thus, our present studies establish that the effects of *ras* p21 on phospholipid metabolism differ qualitatively from those of v-sis/PDGF-2 and suggest that the specific alterations in Ptd Cho and Ptd Etn hydrolysis observed in *ras* transformants could be important in mediating its oncogenic potential.

Received 31 August; accepted 7 October 1987.

1. Bishop, J. M. *Science* **235**, 305–309 (1987).
2. Gilman, A. G. *Cell* **36**, 577–579 (1984).
3. Berridge, M. J. *Biochem. J.* **212**, 849–858 (1983).
4. Nishizuka, Y. *Science* **225**, 1365–1370 (1984).
5. Berridge, M. J., Heslop, J. P., Irvine, R. F. & Brown, K. D. *Biochem. J.* **222**, 195–201 (1984).
6. Lacal, J. C. & Tronick, S. E. in *The Oncogene Handbook* (Elsevier, New York, in the press).
7. Scolnick, E. M., Papageorge, A. G. & Shih, T. Y. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5355–5358 (1984).
8. Lacal, J. C., Srivastava, S. V., Anderson, P. S. & Aaronson, S. A. *Cell* **44**, 609–617, (1986).
9. McGrath, J. P., Capon, D. J., Goeddel, D. V. & Levinson, A. D. *Nature* **310**, 644–649 (1984).
10. Willumsen, B. M., Christensen, A., Hubbert, N. L., Papageorge, A. G. & Lowy, D. R. *Nature* **310**, 583–586 (1984).
11. Paries, G., Hoebel, R. & Roacker, E. *Proc. natn. Acad. Sci. U.S.A.* **84**, 2648–2652 (1987).
12. Wakelam, M. J. O. *et al. Nature* **323**, 173–176 (1986).
13. Doolittle, R. F. *et al. Science* **221**, 275–277 (1983).
14. Grove, R. I. & Schimmel, S. D. *Biochim. biophys. Acta* **711**, 272–280 (1982).
15. Guy, G. R. & Murray, A. W. *Cancer Res.* **42**, 1980–1985 (1982).
16. Bocchino, S. B., Blackmore, P. F. & Exton, J. H. *J. biol. Chem.* **260**, 14201–14207 (1985).
17. Muir, J. G. & Murray, A. W. *J. cell Physiol.* **130**, 382–391 (1987).
18. Irving, H. R., Exton, J. H. *J. biol. Chem.* **262**, 3440–3443 (1987).
19. Stacey, D. W. & Kung, H. *Nature* **310**, 508–511 (1984).
20. Hagag, N., Lacal, J. C., Graber, M., Aaronson, S. A. & Viola, M. V. *Molec. cell. Biol.* **1**, 1984–1988 (1987).
21. Vara, F., Schneider, J. A. & Rozengurt, E. *Proc. natn. Acad. Sci. U.S.A.* **82**, 2384–2388 (1985).
22. Pasti, G., Lacal, J. C., Warren, B. S., Aaronson, S. A. & Blumberg, P. M. *Nature* **324**, 375–377 (1986).
23. Lacal, J. C., Fleming, T. P., Warren, B. S., Blumberg, P. M. & Aaronson, S. A. *Molec. cell. Biol.* (in the press).
24. Bar-Sagi, D. & Feramisco, J. R. *Science* **233**, 1061–1068 (1986).
25. Fleischman, L. F., Chahwala, S. B. & Cantley, L. C. *Science* **231**, 407–410 (1986).
26. Wolfman, A. & Macara, I. *Nature* **325**, 359–361 (1987).
27. Preiss, J. *et al. J. biol. Chem.* **261**, 8597–8600 (1986).
28. Chiarugi, V. P., Pasquali, F., Vannuchi, S. & Ruggiero, M. *Biochem. biophys. Res. Commun.* **141**, 591–599 (1986).

29. Besterman, J. M., Duronio, V. & Cuatrecasas, P. *Proc. natn. Acad. Sci. U.S.A.* **83**, 6785-6789 (1986).
30. Sugimoto, Y., Whitman, M., Cantley, L. C. & Erikson, R. L. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2117-2121 (1984).
31. Berridge, M. J. *Biochem. J.* **212**, 849-858 (1983).
32. Robbins, K. D., Devare, S. G., Reddy, E. P. & Aaronson, S. A. *Science* **218**, 1131-1133 (1982).
33. Robbins, K. C., Leal, F., Pierce, J. H. & Aaronson, S. A. *EMBO J.* **4**, 1783-1792 (1985).
34. Irvine, R. F., Anggard, E. A., Letcher, A. J. & Downes, C. P. *Biochem. J.* **229**, 505-511 (1985).

Elevated *c-myc* expression facilitates the replication of SV40 DNA in human lymphoma cells

Marie Classon, Marie Henriksson, Janos Sümegi, George Klein* & Marie-Louise Hammaskjöld†

Department of Tumor Biology, Karolinska Institute, Box 60 400, S-10401 Stockholm, Sweden

The *v-myc* oncogene can induce tumours in haematopoietic, mesenchymal and epithelial tissues¹. The corresponding *c-myc* proto-oncogene can contribute to the genesis and/or the progression of an equally wide variety of tumours when activated by retroviral insertions, chromosomal translocations or gene amplification². The *c-myc* gene product is a DNA-binding, nuclear phosphoprotein that is involved in the control of cell proliferation and possibly in DNA synthesis². The replication of Simian virus 40 (SV40) is a useful model system to study eukaryotic DNA replication as the virus relies almost entirely on cellular DNA replication apparatus. The SV40-based vector, pSVEpR4, replicates poorly in the human BJAB lymphoma line and in most human cells, but replicates well in Burkitt lymphoma lines, which have fused immunoglobulin and *c-myc* genes, resulting in high *c-myc* expression. Cotransfection of the BJAB cells with a *c-myc*-expressing construct (pI4-P6) increased the replication of pSVEpR4 tenfold. Our findings indicate that overexpression of the *c-myc* gene product allows the replication of SV40 in human lymphoma cells, suggesting that *c-myc* is involved in the control of replication.

We have previously found that transient transfection of the Epstein-Barr virus (EBV)-negative Burkitt lymphoma (BL) cell line RAMOS (ref. 3) with a SV40-based vector, pSVEpR4, which contains the EBNA-1 encoding region of EBV (pBamK), led to unexpectedly high expression of the EBNA-1 protein⁴, indicating that the recombinant vector was replicating efficiently. We further assayed the replication of pSVEpR4 in EBV-carrying and EBV-negative human BL cell lines. Cells were transfected with the pSVEpR4 vector^{4,5}, the plasmid DNA was extracted after 48 hours⁶ and digested with the restriction endonucleases *DpnI*, *MboI* and *BamHI*. The 5.2-kilobase (kb)-long pSVEpR4 vector contains one *BamHI* cleavage site and 12 recognition sites for the isoschizomers *DpnI* and *MboI*. *DpnI* can only cleave the DNA if the adenine residues on both strands have been methylated by the bacterial *dam* (ref. 25) methylase, whereas *MboI* does not cleave DNA with the bacterial methylation pattern. Replicated pSVEpR4 molecules therefore appear as 5.2-kb-long fragments after *DpnI*/*BamHI* double digestion. Plasmid DNA molecules that have not replicated at least once in the human cells remain sensitive to *DpnI* and appear as smaller fragments. As shown in Fig. 1, the pSVEpR4 plasmid DNA replicated well not only in the RAMOS line, as expected, but also in the three other BL cell lines, Raji⁷, Daudi⁸ and Loukes⁹, irrespective of whether they carry EBV genomes or not. These lines give high *c-myc* expression due to the BL-

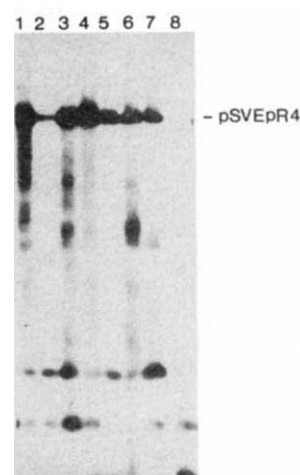


Fig. 1 Southern blot analysis of transfected pSVEpR4 DNA in human lymphoma lines. The cells were transfected with 5 µg of plasmid using DEAE-dextran^{4,5}. Between 5-15% of the cells were transfected as monitored in parallel experiments using pBamK. 48 h post-transfection, plasmid DNA was purified from 10⁶ cells using the procedure of Hirt⁶. The DNA samples from the human lymphoma line RAMOS (lane 1), BJAB (lane 2), BJAB/B95-8 (lane 3), BJAB-HRIK (lane 4), Daudi (lane 5), Loukes (lane 6) and Raji (lanes 7, 8) were digested with *DpnI*/*BamHI*. *BamHI* linearizes pSVEpR4; *DpnI* will only cleave the DNA when it is methylated by the bacterial *dam* methylase²⁴. The sample in lane 8 was digested with *MboI* in addition to *DpnI* and *BamHI*. Digested DNA samples were fractionated by electrophoresis using 0.8% agarose A (Pharmacia), transferred to nitrocellulose filters (0.45 µm Hybond C, Amersham) according to the method of Southern²⁵ and hybridized to ³²P-labelled SV40 DNA. The filters were washed and exposed to X-ray film (XR5 Fuji film) at -70 °C. As the DNA used for transfections was grown in HB 101 (*dam*+ *E. coli*), only DNA molecules that have replicated in the eukaryotic cells will be resistant to digestion by *DpnI* and migrate as unit length DNA. *MboI* will only cleave DNA molecules replicated in the eukaryotic cells.

associated *c-myc*/immunoglobulin translocation¹⁰. Only a small amount of replication was observed, however, in BJAB (ref. 11), an EBV-negative B cell lymphoma line lacking the immunoglobulin/*c-myc* translocation¹², with resulting low expression of *c-myc*¹³. In two EBV-converted sub-lines of BJAB, BJAB/B95-8 (ref. 14) and BJAB-HRIK (ref. 15), pSVEpR4 replicated well. The sub-lines of BJAB express *c-myc* messenger RNA to a degree comparable to the BL lines¹³ and support the replication of pSVEpR4 20-50-fold better than BJAB cells, as measured by densitometry. These results indicated a possible relationship between *c-myc* expression and the ability of pSVEpR4 to replicate in human lymphoma lines.

The importance of *c-myc* encoded protein in promoting the replication of pSVEpR4 was investigated in cotransfection experiments using pSVEpR4 and a *c-myc* expressing vector, pI4-P6. *c-myc* vectors were constructed by inserting a 3.8-kb long *SmaI*-*EcoRI* fragment containing the coding exons of the human *c-myc* gene into the pSVEpR4 and pI4 vectors respectively (Fig. 2a). When the *c-myc* carrying vectors pSVEpR4-P6 and pI4-P6 were transfected into CV-1, RAMOS and BJAB cells the expressed *c-myc* product was a nuclear protein of relative molecular mass (*M_r*) 62,000 (62K), as shown by indirect immunofluorescence (Fig. 2b, c) and immunoprecipitation (Fig. 2d). The vector-encoded *c-myc* protein was phosphorylated like the endogenous form (data not shown).

The result of the cotransfection experiments is the progressive enhancement of pSVEpR4 replication as a function of the amount of pI4-P6, as shown in Fig. 3a, suggesting a direct relation between the degree of *c-myc* expression and the replication of pSVEpR4. The amount of pSVEpR4 replication increased

* To whom correspondence should be addressed.

† Present address: Department of Microbiology, State University of New York at Buffalo, New York 14214, USA.

Fig. 2 Expression of the *c-myc* protein in cells transfected with the *c-myc*-containing vectors pSVEpR4-P6 and pI4-P6. *a*, Construction of *c-myc* vectors. pSVEpR4 contains the entire early region of SV40 expressing the large (T) and small (t) tumour antigens, the SV40 origin of replication and the late promoter sequences⁴. pI4 contains the SV40 origin of replication, the early promoter and the late region encoding the capsid antigens. The vector also contains the *Hind*III-*Bam*HI fragment from the plasmid pR4 (ref. 4). The *Apa*I-*Bcl*I fragment of SV40 containing polyadenylation signals was inserted into the *Cla*I site in pR4 after the addition of *Cla*I linkers. The *Hind*III site 3' of the early promoter was converted into a unique site for the insertion of foreign genes by the addition of *Xba*I linkers. The 3.8-kb-long *Sma*I-*Eco*RI fragment (P6) containing the *c-myc* coding exons from the human genomic clone pMC41HE (ref. 26) was cloned into the pSVEpR4 and pI4 vectors. *b, c*, Indirect immunofluorescence staining using rabbit polyclonal *v-myc* antisera²⁷ of *b*, CV-1 cells and *c*, RAMOS cells 48 h after transfection with pSVEpR4-P6. *d*, Immunoprecipitation of the *c-myc* protein from transfected cells. Lane 1, COLO 320 HSR. This cell line carries an amplified *c-myc* gene and expresses large amounts of *c-myc* protein²⁸. Lanes 2 and 3 show immunoprecipitates from CV-1 cells; lanes 4-7 immunoprecipitates from COS cells. The cells were transfected with pSVEpR4 (lanes 2, 4), pSVEpR4-P6 (lanes 3, 5), pI4 (lane 6) and pI4-P6 (lane 7). *M*, for standard markers are shown to the right (in K).

Methods. $\sim 5 \times 10^6$ cells were labelled with $125 \mu\text{Ci ml}^{-1}$ of [^{35}S]methionine in medium without methionine at 48 h post-transfection. The cells were washed with PBS and lysed with 10 mM Tris-HCl, pH 7.5, 50 mM NaCl, 0.5% Nonidet P-40, 0.5% deoxycholate, 0.5% SDS, supplemented with 2 mM PMSF. The extracts were sheared by passing them through a thin needle and insoluble material was removed by centrifugation and subsequently incubated with α -*v-myc* sera and 100 μl 10% protein A Sepharose for 1 h on ice. The immunoprecipitates were washed three times with RIPA buffer (50 mM Tris-HCl, pH 7.2, 150 mM NaCl, 1% Triton X 100, 1% deoxycholate, 0.1% SDS) and three times with 10 mM Tris-HCl, pH 7.5. The precipitates were resuspended in 100 μl 2 $\times$ sample buffer (80 mM Tris-HCl, pH 6.8, 4% SDS, 20% glycerol, 10% β -mercaptoethanol and bromophenolblue), heated for 10 min at 80 $^{\circ}\text{C}$ and fractionated by 8% SDS-PAGE (ref. 29). The ^{35}S labelled polypeptides were detected by fluorography using Amplify (Amersham).

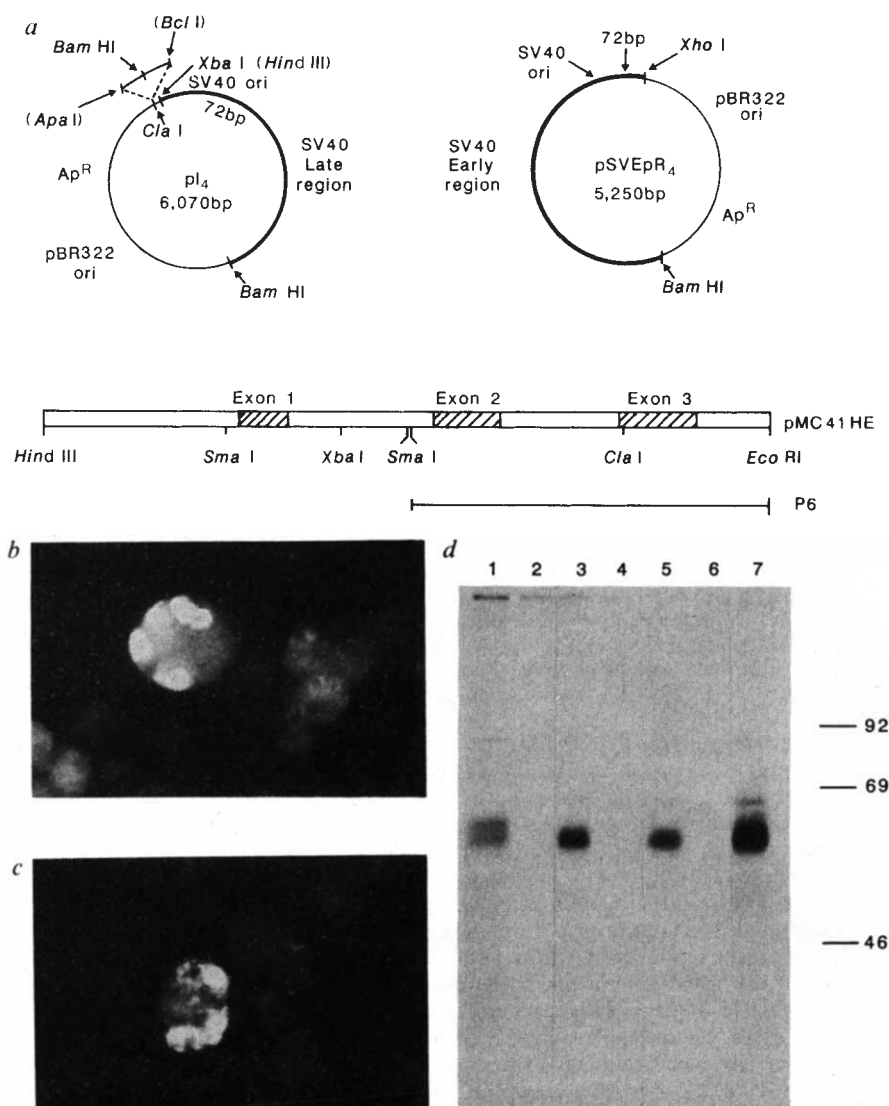


Fig. 3 Effect on the replication of pSVEpR4 of increasing levels of *a*, *c-myc* and *b*, SV40 T antigen. 5 μg (1.4 pmol) of pSVEpR4 vector DNA was transfected with increasing amounts of the *c-myc* expressing pI4-P6 construct (*a*) or the origin-defective pMKS_{SV}_{ori} plasmid (*b*), respectively. At 48 h post-transfection low molecular weight DNA was isolated and analysed as described in Fig. 1. *a*, The total amount of DNA cotransfected with the pSVEpR4 vector was kept constant by addition of pI4 vector, obtaining the following molar ratios of pI4/pI4-P6 DNA: 2.5 pmol (10 μg)/0.0 pmol (lane 1), 1.25 pmol/1.25 pmol (lane 2) and 0.6 pmol/1.9 pmol (12.3 μg , lane 3). *b*, The pMKS_{SV}_{ori} DNA was used in the following concentrations in transfection assays in BJAB (lanes 1-6) and RAMOS (lanes 7-11) cells respectively: 0.04 pmol (lane 1), 0.4 pmol (lane 2), 1.0 pmol (lane 3), 2.1 pmol (lane 4), 4.3 pmol (15 μg , lanes 5, 6), 0.0 pmol (lane 7), 0.7 pmol (lane 8), 1.4 pmol (lane 9) and 2.8 pmol (lanes 10, 11). The samples in lanes 6 and 11 were digested with *Mbo*I in addition to *Dpn*I and *Bam*HI.

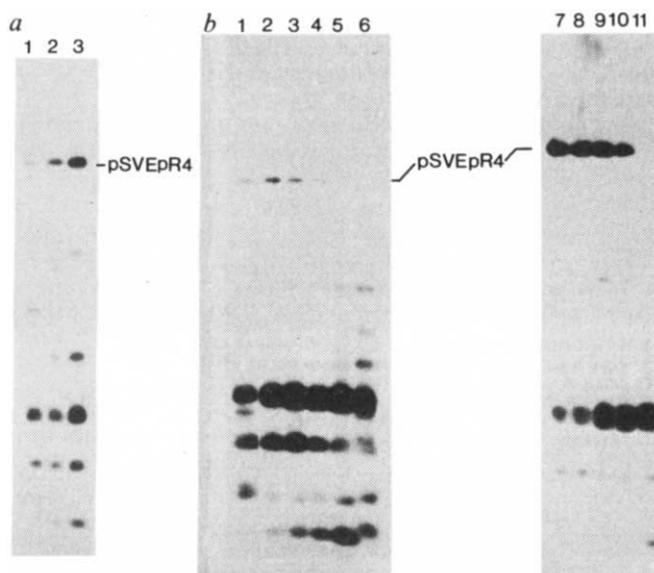
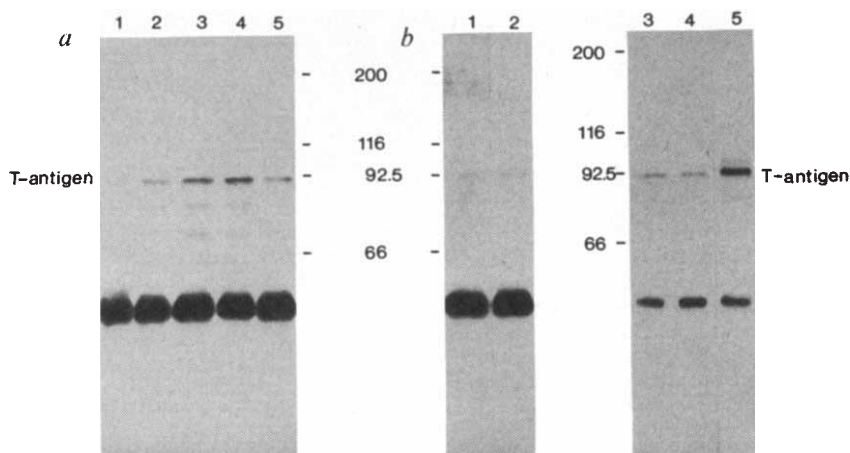


Fig. 4 Immunoblot of SV40 large T antigen immunoprecipitated from transfected cells. *a*, BJAB cells transfected with 5 μ g (1.4 pmol) pSVEpR4 and 0.7 pmol (lane 1), 1.4 pmol (lane 2), 2.8 pmol (lane 3) or 4.2 pmol (15 μ g, lane 4) pMKSV_{ori} plasmid. 5×10^6 CV-1 cells transfected with 1.4 pmol pMKSV_{ori} (lane 5). *b*, BJAB (lanes 1, 2) and RAMOS (lanes 3, 4) transfected with 1.4 pmol (5 μ g) of pMKSV_{ori} and 2.5 pmol (10 μ g) pI4 (lanes 1, 3) or 0.6 pmol pI4/1.9 pmol (12.3 μ g) pI4-P6 (lanes 2, 4). 5×10^6 CV-1 cells transfected with 1.4 pmol (5 μ g) of pSVEpR4 (lane 5). M_r for standard markers are shown (K) to the left.

Methods. 4×10^7 cells were lysed in 20 mM Tris-HCl, pH 7.5, 150 mM NaCl, 5 mM EDTA, 0.5% Triton X-100, 0.5% deoxycholate, 0.1% SDS and 2 mM PMSF. The extracts were sonicated (Soniprep 150, MSE) and insoluble material was removed by centrifugation. The resulting supernatants were incubated with T antigen monoclonal Pab 416 and 20 μ l 50% Protein A Sepharose overnight at 4 °C. The immunoprecipitates were washed twice in lysis buffer, once in lysis buffer/PBS and then finally in PBS before adding 100 μ l 2 $\times$ sample buffer. The samples were heated for 10 min at 100 °C and then fractionated by 7.5% SDS-PAGE (ref. 29). Each lane was loaded with samples equal to 10^7 cells. Electrophoretic transfer to nitrocellulose was carried out according to Towbin³⁰. After preabsorption in PBS supplemented with 0.5% non-fat dry milk (PBSM), the filters were incubated with the T antigen monoclonal in PBSM overnight at 4 °C. Following washes in PBSM the filters were incubated with alkaline phosphatase-conjugated rabbit anti-mouse IgG (Sigma) in PBSM. After washing in PBS-0.1% Tween 20 the filters were stained with 50 mM Tris-HCl, pH 8.8, 2 mM MgCl₂, 0.2 mg ml⁻¹ alpha-naphthylphosphate (Sigma) and 0.5 mg ml⁻¹ fast red salt (Sigma).



tenfold in the cotransfected BJAB cells, as measured by densitometry. Enhanced replication was also seen in RAMOS cells, although these have an elevated endogenous *c-myc* expression (data not shown). Replication of SV40 in permissive monkey cells requires replication factors provided by the host cell, in addition to the virally encoded T antigen. T antigen is a limiting factor in permissive cells, where it determines the extent of viral DNA replication¹⁶, but in non-permissive murine cells even large amounts of wild-type SV40 T antigen do not overcome the non-permissive environment for SV40 replication¹⁷. The supply of T antigen could possibly be limiting in BJAB cells: if so, the *c-myc* protein could mediate its effect through an increase in T antigen expression from the transfected pSVEpR4 DNA. To test this possibility, we transfected BJAB and RAMOS cells with pSVEpR4 and an increasing amount of an origin-defective, T antigen-expressing plasmid, pMKSV_{ori} (ref. 18). The transfected cells had increased expression of T antigen (Fig. 4a), although this was not sufficient to significantly increase the amount of replicated pSVEpR4 DNA (Fig. 3b). No enhanced expression of T antigen was detected from the pMKSV_{ori} plasmids when they were cotransfected into BJAB and RAMOS cells with increasing amounts of pI4-P6 (Fig. 4b). This indicates that even large amounts of *c-myc* protein fail to stimulate gene expression from the early SV40 promoter, in agreement with another report¹⁹.

The normal restriction of human cells to be semi-permissive for SV40 growth is not observed in two human embryonic kidney

(HEK) cell lines transformed by the adenovirus early genes, suggesting that the E1 proteins can generate a permissive environment^{20,21}. This effect of adenovirus E1 proteins is similar to that shown here, where high intracellular concentrations of the *c-myc* protein render human cells permissive for SV40 replication. Although the replication of pSVEpR4 in human lymphoid cell lines requires T antigen, the *myc*-dependent enhancement of replication was not a result of an increased SV40 T antigen level. However, we cannot rule out the possibility that the *c-myc* encoded protein could affect the modification of the T antigen, or could change the interaction of T antigen with cellular factors.

The complete mechanism by which the *c-myc* and E1 proteins create a permissive environment to which the SV40 replicon responds is not yet known, but it is tempting to speculate that they act through a common pathway, belonging as they do to the same group of "immortalizing" oncogenes^{22,23}.

We thank K. Alitalo for antisera to *v-myc* protein and procedures for *myc* protein analysis, D. Lane for the Pab 416 T antigen monoclonal and hybridoma cells, H. Ozer for the pMKSV_{ori} plasmid, B. Ehlin-Henriksson for technical assistance, K. Swiech for help with fluorescence photography and drawings, and Svartvita Kopior for photography. This work was supported by the National Cancer Institute and by the Swedish Cancer Society. M.-L.H. and J.S. were recipients of fellowships awarded by the Cancer Research Institute Inc. and by Concern Foundation.

Received 18 June; accepted 8 October 1987.

- Moscovici, C. & Purchio, A. F. *Adv. Viral Oncol.* **1**, 83-106 (1982).
- Bishop, J. M. *Science* **235**, 305-311 (1987).
- Klein, G., Giovannella, B., Westman, A., Stehlin, J. S. & Mumford, D. *Intervirology* **5**, 319-334 (1975).
- Hammarskjöld, M.-L., Wang, S.-C. & Klein, G. *Gene* **43**, 41-51 (1986).
- McCutchan, J. H. & Pagano, J. J. *natn. Cancer Inst.* **41**, 351-357 (1968).
- Hirt, B. *J. molec. Biol.* **26**, 365-369 (1967).
- Pulvertaft, R. I. V. *J. clin. Path.* **18**, 261-273 (1965).
- Klein, E. *et al. Cancer Res.* **28**, 1300-1310 (1968).
- Van Santen, V., Cheung, A. & Kieff, E. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1930-1934 (1981).
- Klein, G. *Cell* **32**, 311-315 (1983).
- Menezes, J., Leibold, W., Klein, G. & Clements, G. B. *Biomedicine* **22**, 276-284 (1975).
- Klein, G. *et al. Int. J. Cancer* **18**, 639-652 (1976).
- Wennborg, A. *et al. Int. J. Cancer* **40**, 202-206 (1987).
- Clements, G. B., Klein, G. & Povey, S. *Int. J. Cancer* **16**, 125-133 (1975).
- Fresen, K. O. & zur Hausen, H. *Int. J. Cancer* **17**, 161-166 (1976).

- Gerhard, R. D., Guggenheimer, R. A. & Gluzman, Y. *J. Virol.* **61**, 851-857 (1987).
- Gerhard, R. D. & Gluzman, Y. *Molec. cell. Biol.* **5**, 3231-3240 (1985).
- Gluzman, Y., Sambrook, J. & Frisque, J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3398-3902 (1980).
- Kaddurah-Daouk, R., Greene, J. M., Baldwin, A. S. Jr & Kingston, R. E. *Genes Dev.* **1**, 347-357 (1987).
- Lewis, E. D. & Manley, J. L. *Nature* **317**, 172-175 (1985).
- Lebkowski, J. S., Clancy, S. & Calos, M. P. *Nature* **317**, 169-171 (1985).
- Land, H., Parada, L. F. & Weinberg, R. A. *Science* **234**, 467-470 (1983).
- Ruley, H. E. *Nature* **304**, 602-606 (1983).
- Geier, G. E. & Modrich, P. *J. biol. Chem.* **254**, 1408-1413 (1979).
- Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
- Watt, R. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 6307-6311 (1983).
- Winqvist, R., Saksela, K. & Alitalo, K. *EMBO J.* **12**, 2947-2950 (1984).
- Alitalo, K., Schwab, M., Lin, C. C., Varmus, H. & Bishop, J. M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 1707-1711 (1983).
- Laemmli, U. K. *Nature* **227**, 680-685 (1970).
- Towbin, H. & Gordon, J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4350-4355 (1979).

The treason of the clerks

Conor Cruise O'Brien

Academic Freedom and Apartheid: The Story of the World Archaeological Congress.

By Peter Ucko. Duckworth, London: 1987. Pp.305. Hbk £18; pbk £9.95.

PETER Ucko is Professor of Archaeology at Southampton University and was appointed National Secretary of the British Congress of the International Union of Pre- and Protohistoric Sciences (IUPPS). The Congress was due to be held at Southampton in September 1986. In September 1985, however, the British Executive, without consultation with the IUPPS, decided to ban the participation of South African scholars. In January 1986, the International Executive Committee of the IUPPS declared that it "cannot accept any meeting where participation is subjected to non-scientific considerations as being organized under the auspices of the Union. Therefore the Executive Committee refuses to recognize the Southampton meeting as the IUPPS Congress". The British National Committee went ahead with the Southampton conference, while continuing to describe it as the "World Archaeological Conference". The XIth IUPPS Congress, originally scheduled for Southampton, was instead held at Mainz in September of this year.

The transactions discussed in *Academic Freedom and Apartheid* had no effect whatever on apartheid, as far as I know. But they did have a significant and negative effect on respect for academic freedom, and therefore on academic freedom itself, both in Britain and in South Africa, and to some extent elsewhere. The events that led up to the Southampton Congress, and the split in the IUPPS, therefore form a not insignificant chapter in the intellectual history of the late twentieth century.

As British National Secretary, Professor Ucko was at the centre of these events. The publication of his book about them is therefore to be welcomed, though it should be taken with more than one grain of salt. It is, perhaps inevitably, a strongly partisan account. Even Neal Ascherson, in a distressingly gushing foreword, gags a bit towards the end, and concedes: "Not all arguments here are scrupulous or fair".

In retrospect, the decision to ban the South African scholars is made to appear high-minded and idealistic; motivated by sincere abhorrence for apartheid (which in fact several of the scholars banned also abhor). But the documents contemporary with the relevant transactions, cited by Professor Ucko, reflect a different motivation to that claimed in retrospect. The following extract is from the minutes of the British Executive Committee of 15



Peter Ucko — principle or pragmatism?

October 1985, confirming the decision to ban the South African scholars:

The Treasurer reiterated his view, which the Executive Committee accepted, that had South African attendance not been prevented for the WAC 1986 then the financial viability of the whole Congress would have been likely to have collapsed. It was with this advice in mind that the Executive, mindful of its duties as Directors ... reiterated its view that, on purely pragmatic grounds, the Executive had no other option than to repeat its previous decision to ban South African participation from the Congress.

Subsequent statements — not only by Professor Ucko but also by some of his opponents — suggest a painful clash of principle: essentially, the principle of racial freedom, on the one side, against the principle of academic freedom on the

other. Very tragic and noble, if true. But what the record shows is not a clash between one principle and another, but pragmatism winning hands down over any consideration of principle whatever.

Putting it bluntly, what happened was that a bunch of British pressure groups put the heat on, and that this group of British academics caved in. The pressure groups concerned included Southampton City Council, Southampton University itself, the National Union of Students and the Association of University Teachers. These could probably have made it impossible for the Congress to be held in Southampton; a similar line-up might well have made a Congress impossible anywhere in Britain. But the disquieting thing is that an Executive, composed of distinguished British scholars, should have

unanimously decided, when it came to the crunch, that it was more important to hold a particular Congress than to stick to the principles of the universality of science and scholarship, and the free exchange of ideas.

It would be too much to say that these principles were thrown to the wind. They were just quietly discarded, amid muffled noises about the Treasurer's report.

Later indeed, when the IUPPS made its attitude clear, some of the Britons concerned began to give more thought to the principles they had just dropped. Even Peter Ucko mourned for a time over the breach of academic freedom. In a letter to one of the banned South African scholars, he wrote: "We do indeed keenly regret the breach of the principles of free academic interchange

implicit in our decision".

As time went on, the regret diminished and Professor Ucko now writes of academic freedom with open contempt. In his Introduction he writes: "to discuss the issue of academic free speech is almost an obscenity in the context of South Africa". And the last sentence of his book runs: "It is my hope that the story of the World Archaeological Congress may help them [scholars] to resolve this dilemma, to come out from behind the shaky edifice of their own academic freedom, and turn their attention to the issue of freedom itself".

But academic freedom is a functional part of freedom in general. Academic freedom is a set of conventions built up to protect the freedom to teach and to learn;



Educated response — unrest in South Africa itself, as students at the University of Witwatersrand commemorate the Sharpeville massacre.

the freedom to publish and to exchange information. If those freedoms are to be dismissed as bourgeois luxuries, as Professor Ucko's rhetoric implies, it seems doubtful whether what remains can reasonably be described as freedom at all.

It is easy to protest one's devotion to freedom-at-large. It may be a little more difficult to defend a particular set of freedoms, in relation to which you have responsibility, when those freedoms come under pressure. Academic freedom may indeed be "a shaky edifice", as Professor Ucko says. But those responsible for "a breach of the principle of free academic exchange" — a breach acknowledged by Professor Ucko himself — hardly have the right to reproach the edifice so breached for its shaky condition.

If it could be demonstrated that restrictions on academic freedom would relieve the oppression of blacks in South Africa, then there would indeed be a case for considering such restrictions. But no such demonstration has been made. As far as Pretoria is concerned, the so-called "academic boycott" is a joke. The South African institutions damaged by the "boycott" — or rather by the outbreaks of student violence organized in the name of the boycott — are the liberal English-speaking universities, those which alone take black students in significant numbers. To suppose that Pretoria feels adversely affected by damage done to those institutions is like supposing that Mrs Thatcher could be brought to her knees by an international boycott of the North London Polytechnic.

Professor Ucko does not attempt to show that the "academic boycott" or the "World Archaeological Congress" has damaged the Pretoria regime, or is capable of damaging it. What he does is to say repeatedly (and quite truly) that the African National Congress, UNESCO, the UN Special Committee on Apartheid and a long et cetera, reaching down to the

Chairman of the Southampton City Council, all demand that South African scholars be kicked out of international scholarly gatherings. That scholars should assert any will or mind of their own in scholarly matters in the presence of such a formidable array of non-scholarly authorities seems incomprehensible to Professor Ucko. He points out, for example that "insistence of South African archaeologists ... to attend international gatherings ... goes directly contrary to the A.N.C. position". That apparently should settle the matter, for any archaeologist worthy of his salt.

The idea of the verdict of an institution as automatically binding on individual intellects, is an old one, but one that seems to be on its way back. For Professor Ucko and his friends, "individualists" are people who have to be made to toe the line, in one way or another. Thus, when the Southampton lobby were trying to swing votes to their side on the eve of a critical IUPPS meeting in Paris, they went to work on (among others) a certain Mexican prehistorian, Professor J. Lorenzo. Mexico — meaning the Mexican government — was sound on the boycott, from the lobby's point of view. But might this particular Mexican scholar have the nerve to vote differently from what his government wanted? It seemed he might. "We were worried by reports that Professor Lorenzo was an individualist." The Southampton lobby could not win him round, so they put the pressure on to keep him away from the meeting. "In the event", says Professor Ucko of his Mexican colleague, "he also did not materialize in Paris — maybe he had been convinced by a letter from the A.N.C. or, as a result of a number of telephone calls to the British Ambassador in Mexico, by a word from his Foreign Minister!" (Professor Ucko's exclamation mark). Similar pressure seems to have been applied to an Indian archaeologist whose stance is

described as "equivocal".

It is well known that in most Third World countries scholars are, for various reasons, much more vulnerable to governmental pressure than their counterparts in the West. But that Western scholars could deliberately use that vulnerability, in order to force their Third World colleagues into line, through signals to their governments, is something that I would not have believed did I not have Professor Ucko's gleeful testimony to that effect (on p. 95 of this book).

Julian Benda wrote of "the treason of the clerks". He meant by that, not treason to a particular country or ideology, but the betrayal by intellectuals of their actual function. Intellectuals betray their function, in Benda's view, when they subordinate it to the demands of an ideology; when they subordinate individual judgement to the demands of a political collective; and when they seek to coerce other intellectuals into similar patterns of behaviour.

Anyone interested in the state of "the treason of the clerks" in Britain today should read *Academic Freedom and Apartheid*. And anyone reading it should remember that Professor Ucko is not some lonely, marginal figure. It is not enough to say that the Association of University Teachers are on Professor Ucko's side; the AUT were among the bodies who pushed this particular teacher towards the position which he has now assumed. So, from the Benda point of view — which I share — the rot has gone pretty far in these islands. And if we may judge from the Southampton story, it has gone a lot further in Britain at this time than in the academies of France, Germany or North America. □

Conor Cruise O'Brien, *Whitewater*, *Howth Summit*, Dublin, Ireland, is Pro-Chancellor of the University of Dublin. He was Vice-Chancellor of the University of Ghana from 1962 to 1965, and Editor-in-Chief of *The Observer* newspaper from 1979 to 1981.

Skeletons in the cupboard

C.K. Brain

Bones of Contention: Controversies in the Search for Human Origins. By Roger Lewin. *Simon & Schuster: 1987. Pp. 348. \$19.95.*

"THERE are only two kinds of science", a colleague remarked to me recently, "good science and bad science". What, one wonders, are the hallmarks of each? Surely, objectivity — the drawing of conclusions unmodulated by conscious or unconscious bias — would be a criterion of "good" science. But when such a criterion is applied to palaeoanthropology, as is so compellingly done by Roger Lewin in his new book, one is forced to concede that this discipline still has a greater than usual amount of evolving to do on its road to objectivity.

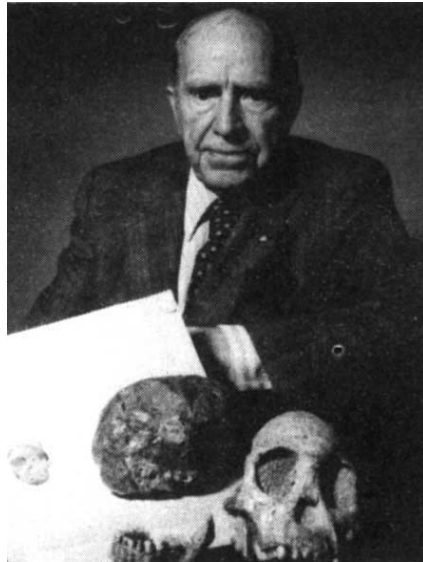
Palaeoanthropology is a popular subject, with many people eager to know something of their origins and the course that their ancestors followed during their emergence from the animal world. Roger Lewin, who graduated in biochemistry from Liverpool University, worked as an editor at *New Scientist* and is currently editor of research news at *Science*, has been caught up in the excitement of palaeoanthropology for some time. He is often seen at international meetings concerned with early man and has co-authored two best-selling books, *Origins* and *People of the Lake* with Richard Leakey.

His present book is about "controversies in the search for human origins" and, in presenting these, Lewin has made use of private letters, published papers and interviews with the personalities involved. He reports on events which, at the time, generated a good deal of acrimony, yet he does so with impartiality and even-handedness to expose the roots of each dispute. The result is an enthralling narrative of people catapulted into prominence by their fossil finds and of their passionate attitudes towards their subject. But Lewin goes beyond personalities: he documents the state of a science in transition as its proponents struggle to dissociate their subconscious and emotional responses from the dictates of their intellects.

The controversies Lewin has chosen to investigate have been major issues in the annals of palaeoanthropology. He deals at length with the rejection of the Taung child by the anthropological community of the day because it did not conform to what an ancestral hominid was expected to have looked like, nor had it lived on the right continent. He writes of the initial acceptance of *Ramapithecus* as the earliest of all hominids and then

describes the motives behind its dismissal. He discusses the preoccupation of Louis and Richard Leakey with their search for the earliest representative of the genus *Homo*, and how this obsession eventually led to a remarkable dispute about the inferred ages of fossil sequences from the Omo Valley and Koobi Fora, centring on the antiquity of the 1470 *Homo* skull and on the KBS Tuff which was its age marker. Finally, Lewin documents the extraordinary events leading to the naming of Lucy and her clan, and the furore which the designation *Australopithecus afarensis* unleashed.

Palaeoanthropology is not the only branch of science to generate controversy, but it is remarkable for the amount it generates and for the intensity of emotional involvement that is shown by its



Cause of controversy — R.A. Dart with the Taung skull.

participants. It is this emotional involvement that leads to a breakdown in objectivity and a deterioration in the quality of the science. Why should emotions be so much in evidence? The underlying reasons that emerge from Lewin's book are human frailty at the individual level and community influences that operate subconsciously on the investigator.

The human frailties manifested by palaeoanthropologists are surely no different to those to be found among researchers in other branches of science, but their focus is greatly sharpened. The reason is that palaeoanthropology is a highly visible discipline where a single discovery can make its finder famous overnight. The stakes are high but spectacular fossils are hard to find: Louis and Mary Leakey searched East Africa for over 30 years before the first really important hominid skull came into their hands. So it is understandable that finders of significant fossils should wish to make the most of them, to ensure that they are seen as being on the direct lineage to modern man and are the oldest representatives of this

prestigious line ever discovered. Furthermore, there is a strong temptation to give the fossil a new name to which that of the describer will be linked in perpetuity.

Such expectations are understandable and provide a spur and motivation during the exhausting and unproductive labour which generally precedes the exciting discoveries. But they do mean that the finder becomes emotionally involved with the fossil which in turn emerges as an extension of the finder's ego, sensitive and vulnerable. The outcome is predictable and examples of it may be perceived throughout the history of palaeoanthropology.

But there are other and more subtle influences that affect the objectivity of palaeoanthropologists, and Lewin draws attention to them in the book. He airs some little-known conclusions reached recently by Misia Landau, who sees many writers on human evolution as storytellers, subconsciously moulding their information into the format of the hero myth, so universally present in folk tales. Elements of the evolution myth are these: the ancestral hominid, our humble hero, starts in the stable environment of the evergreen forest; climatic change expels him from this security and he is forced to embark on a hazardous journey during which he overcomes many trials by developing bipedalism and intelligence; his ultimate triumph is the achievement of humanity. "There is a final irony", says Landau:

Again and again we hear how a hero, having accomplished great deeds, succumbs to pride or hubris and is destroyed. In many narratives of human evolution there is a similar sense that man may be doomed, that although civilization evolved as a means of protecting man from nature, it is now his greatest threat.

Are we, in our writings, unwittingly constrained by such universal narratives that rise from some dim, collective unconscious?

Finally, Lewin reminds us that the way palaeoanthropologists write about their subject is inevitably influenced by humanity's image of itself at that particular moment. Such communal human images, like fashions, change with time. Once it was fashionable to see our earliest ancestors as bloodthirsty killers, now it is not. The way we write tells us as much about ourselves as about our fossil subjects.

Lewin ends his book with the prediction that, in the study of human origins, there will always be bones of contention. But my impression is that, as a result of this exposé of the human frailties eroding the discipline of palaeoanthropology, a new trend to objectivity may yet be set in motion. Roger Lewin may have contributed more to the evolution of a science than he had dared to hope. □

C. K. Brain is Director of the Transvaal Museum, PO Box 413, Pretoria 0001, South Africa.

Back to the drawing board

Anthony F. Aveni

Lines to the Mountain Gods: Nazca and the Mysteries of Peru. By Evan Hadingham. Random House, New York/Harrap, London: 1987. Pp.307. \$22.50, £12.95.

The Mystery of the Nasca Lines. By Tony Morrison. Nonesuch Expeditions, 48 Station Road, Woodbridge, Suffolk, UK: 1987. Pp.154. £14.95.

THERE is nothing quite like them. They have been called terrestrial runways for extraterrestrial astronauts, water trenches for irrigation, Olympic running tracks, gigantic maps of a lost empire, and effigies designed to be seen from above either by the ancient balloonists who constructed them or by their celestial gods.

They are, of course, the Nazca lines of south coastal Peru — figures of birds and other biomorphs of decametric dimensions, star-like patterns from which dozens of lines radiate like spokes from a wheel, geometric drawings of labyrinthine spirals and truncated trapezoids, and long, dead-straight lines (over 1,000 kilometres of them if laid end to end), all made by the process of scraping the top soil off the surface of the pampas. Who built the lines? When were they built? And is there encoded in them some sophisticated knowledge that might be a reflection of ancient intellectual customs?

In this age of scientific sensationalism, we are in need of dependable reading matter about enigmas such as this. Happily, both of these books are among the more responsible accounts of Nazca, but for quite different reasons.

Evan Hadingham was an observer-participant during one of my own field seasons on the Nazca plain, and is a well-informed and critical purveyor of the essential elements of his subject. He is concerned principally with placing the Nazca lines in a broad cultural perspective. A pan-Andean outlook is provided in chapters on such subjects as the brightly coloured ceramics for which the Nazca culture (c. 100 BC–AD 600) is justly famous, and on the irrigation technology that included the building of canals and underground stone-lined aqueducts. Hadingham also reviews Nazca ecology and contemporary indigenous astronomical and calendrical concepts. As there is no written record from which we could ever hope to retrieve this sort of information from the past, we are forced, with reservations, to employ ethnographic analogy.

Tony Morrison is also British and is the reporter who some ten years ago produced a neat little popular text on the lines

(*Pathways to the Gods*). In this new book he is concerned more with chronicling Nazca research by concentrating upon the personalities who have hauled their disciplinary wares and predilections out onto the pampas in attempts to answer the question "Why are the lines there?". These individuals constitute a human zoo that Morrison presumes his readers will find even more interesting than the menagerie of figures — condor, monkey, spider and killer whale — that were etched on the desert over a millennium ago.

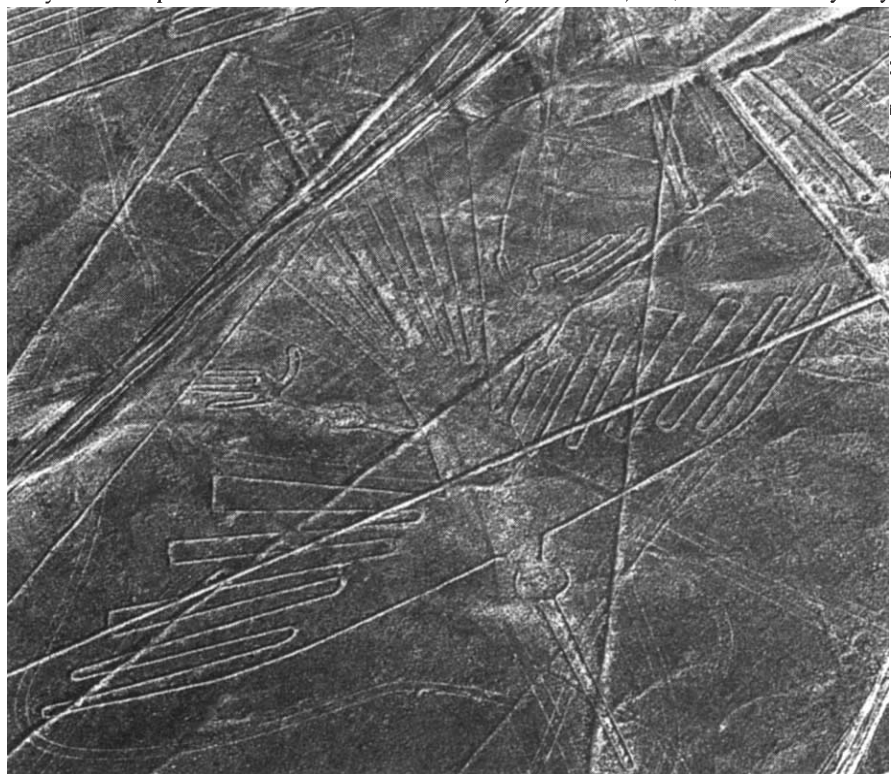
Most famous of them by far is Maria Reiche, a German teacher of mathematics who emigrated from Nazi Germany during the Second World War, and who has spent much of the past half century alone, with little support, often using only pencil and paper to study the lines. She has never wavered from her hypothesis that the vast blackboard at the edge of modern Nazca town is the largest astronomy book in the world, maintaining that the lines were laid out for calendrical purposes to point to celestial objects at the horizon, and that the delicate contours of every figure incorporate one or several precise measuring units.

This rationale has often been applied to large structures built by ancient peoples. Unfortunately, Maria Reiche's unsystematic approach to collecting and analysing data is like that of so many of her successors — Nazca has been a den of undisciplined, disorderly research. Maria is, nevertheless, something of a pioneer for having called attention to the lines, and has selflessly used the profits from the sale of her

booklet *Mystery on the Desert* to hire guards to patrol the pan-American highway adjacent to the figures. Thus clumsy tourists, especially those interested in chatting with extraterrestrials, are prevented from defacing the lines.

Both of these books, particularly Morrison's, treat Reiche as a figure whose ultimate significance to Andean studies is inflated far out of proportion; no fewer than 44 pictures of "Santa Maria" (as the locals call her) adorn these volumes. Morrison, though, does more justice to the native Peruvian researchers who have produced some interesting explanations for the Nazca lines.

Both authors discuss some of the recent research on and about Nazca: Reinhard's study of the great importance of mountain worship and Clarkson's archaeological work on the artefacts on the pampas surface, which imply that the drawings and the long straight lines are of quite different dates. Now we begin to think of the pampas as the unerased blackboard that reveals the activity of several separate episodes of ritual "classroom activity". Morrison does a passable and Hadingham a creditable job of explaining my own research programmes undertaken on the pampas between 1979 and 1984 with Clarkson, Urton and Zuidema. We believe that the straight features are actually an interconnected network of radiating lines that go all the way across the pampas from one river valley to the next, and that they were physically walked upon (lines as ritual pathways is an old idea). We think, too, that astronomy may



Big ideas — a bird on the Nazca plain, 450 feet from beak to tail. The line across its wingspan points from the direction of sunrise in June to that of sunset in December.

Courtesy Tony Morrison

be at least a part of the picture because a marginally significant proportion of straight lines seem to be directed towards key solar positions at the horizon that could have signalled the season when water ran in the canals; this time generally coincides quite closely with the overhead passage of the Sun.

For pictures of the lines, as opposed to people who have studied them, Morrison's book is the preferable of the two with its beautifully illustrated Chapter 6 entitled "The Nazca Lines — A Pictorial Survey". Hadingham has the better bibliography, which will carry lay readers about as far as they should care to go. But in the final analysis both of these books are about more than Nazca and its mysteries, and ought to cause us to reflect on some

deeper questions about ourselves and how we view the past. They also show how much Nazca has become the refuge of the iconoclast. There have been too many short-term programmes and not enough long-term cooperative research, yet the absence of facts fails to deter code-crackers from offering answers. Experts on measuring units, astronomy and computers will not solve this problem. As both authors imply, to understand Nazca we must know about people — and they are far more complicated than higher mathematics. □

Anthony F. Aveni is Charles A. Dana Professor of Astronomy and Anthropology at Colgate University, Hamilton, New York 13346, USA. He is the editor of The Lines of Nazca, to be published next year by the American Philosophical Society.

Before the invasion

Nick Saunders

The Great Journey: The Peopling of Ancient America. By Brian M. Fagan. *Thames & Hudson, London/W.W. Norton, New York: 1987. Pp. 288. £14.95, \$19.95.*

EXPLAINING the presence of America's indigenous population has exercised the minds of scholar and layman alike ever since Columbus accidentally discovered the New World in 1492, promptly but erroneously labelling its inhabitants as Indians. Europeans, their intellectual curiosity sparked by the appearance of 50 Brazilian Indians at the court of King Henri II of France in 1550, offered such explanations as Atlantic crossings by Carthaginians or Phoenicians, and the wanderings of the ten tribes of Israel. In 1589, the Jesuit José de Acosta suggested with some prescience that American Indians had first come from Asia by land and perhaps a short sea voyage. Today, although the arguments are less speculative and more informed, they are no less heated.

American archaeology is divided into two camps by the evidence for early man in the New World. There are those, such as Alan Bryan and Richard MacNeish, who believe that, despite the sparse and often ambiguous data, America was colonized well before c. 10,000 years BP, during the Wisconsin glaciation. In support of this view they cite dates between 30,000 and 19,000 BP from sites such as Tlapacoya in Mexico, Pikimachay in Peru and Boqueirão de Pedra Furada in Brazil. The majority opinion, however, regards human occupation as beginning with the Clovis culture, securely dated to c. 11,000 BP; that culture was replaced a thousand years later by the equally well documented Folsom people. Each is characterized by the distinctively fluted projectile points



New arrival — Explorer René de Laudonnière with Chief Athore in Florida, 1564. Painted by Jacques Le Moyne.

which the peoples fashioned. In his book Brian Fagan fleshes out the picture of the two cultures by graphically reconstructing their hunting and butchering techniques of both mammoth and bison.

The heart of the disagreement between the two schools of thought often lies with the question of whether stone and bone materials were purposely made artefacts or naturally weathered objects. Highlighting the interpretive disagreements is a site with no stone tools at all. Old Crow in the Canadian Yukon has a confused geology of redeposited sands and gravels, and has yielded mammoth bones characterized by distinctive spiral fractures which the excavators interpreted as man made. Together with an obviously manufactured bone flesher, these remains were dated to c. 27,000 BP. Sceptics, who considered the fractures to be the result of freeze-thaw action and subsurface soil movements, and the flesher as a recent but redeposited implement, have had their suspicions vindicated in recent years by a re-dating

of the flesher to c. 1,300 BP.

Similar suspicions surround the site of Calico in southern California, currently dated to between 50,000 and 200,000 years ago. Such early dates suggest a pre-Neanderthal colonization for which there is no corroborative evidence elsewhere in the Americas. The argument for an early Ice Age colonization is difficult to prove archaeologically, but evidence is mounting that slowly pushes back the commonly accepted frontier date of 11,000 BP. Sites such as Meadowcroft in the United States, Bluefish Caves in north-west Canada and Monte Verde in Chile indicate human occupation by at least 12,000 BP.

Proponents of an early arrival take great heart from the geological evidence for a land-bridge, called Beringia, which at various times in prehistory linked the Asian and American landmasses. Soundings, deep-sea cores and palynological evidence suggest a series of alternating sea levels between Alaska and Siberia between 25,000 and 14,000 BP, during the so-called "Duvanny Yar Interval". At this time the sea level stood between 90 and 100 m lower than at present, exposing a wide land-corridor across which early man could have migrated. Fagan judiciously reviews the varying opinions about what type of climate existed on this land-bridge — some argue for a harsh tundra environment inimical to animal life; others believe it to have been a unique ecosystem capable, in places, of supporting both animals and human beings.

Although the disputes on dating are unlikely to be easily resolved, there is more agreement over the distinctive physical characteristics of America's indigenous inhabitants. In 1907 Ales Hrdlicka first proposed, on the basis of physical anthropology, that America was probably populated by immigrants from Asia. In recent years Christy Turner, a specialist in human teeth, has isolated a distinctive set of characteristics such as shovelling of the incisors and number of tooth-roots, which he calls "Sinodonty", and which only occur amongst the indigenous populations of northern Asia and the Americas.

As a non-specialist, Fagan has played safe and has opted for the majority view of a later rather than an earlier occupation of the American continent — though sensibly his final judgement is reserved. *The Great Journey* is an eminently readable account of its subject, in which Fagan carefully pieces together the jigsaw of ambiguous evidence, complex arguments, personalities and discoveries. This is a valuable and long overdue contribution to the popular literature on American archaeology — in addition, it is a fascinating story convincingly told. □

Nick Saunders is Course Tutor in New World Archaeology in the Department of Adult Education, University of Southampton, Southampton SO9 5NH, UK.

Turning a phrase to advantage

Alan Brafield

An Urchin in the Storm: Essays About Books and Ideas. By Stephen Jay Gould. W.W. Norton: 1987. Pp. 255. \$18.95.

A RECENT review in these columns of a book entitled *Communicating Science to the Public* pointed out that the symposium it reflected had contained no contributions from the "great popular communicators" or, indeed, from any scientists. The reviewer also quoted Rutherford's dictum that "no theory was to be taken seriously that could not be explained to a barmaid". Whether Rutherford meant this as a slur on one of my favourite groups of people or as a compliment is a matter for argument, but there is no doubt that popularizing science effectively is as difficult as it is necessary. Gould, one of the very best popularizers, has now produced the latest of his more or less biennial collections of essays. This one is different in that it is a collection of book reviews, but the format and often the content are in the familiar mould.

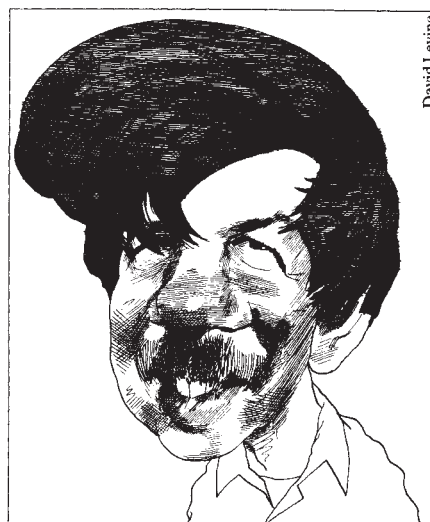
In the preface, Gould maintains that book reviews need not be ephemeral, and contrasts the pedantic and parochial account of a work's content and merit with the type of review in which the writer uses the book's main thrust as an anchor for discussing an issue of wider scope. Naturally he prefers the latter, and says that he cherishes a perceptive essay on embryology and evolution by J.Z. Young although it only mentioned at the end of his own book (*Ontogeny and Phylogeny*) which was the reason for writing the piece. *An Urchin in the Storm* contains 18 essays, written over the past decade, all but one of which originally appeared in *The New York Review of Books*. Gould writes: "We must believe that books, and their ideas, have an enduring meaning worth consideration after an actual title goes out of print". (In fact the latest edition of *Books in Print* shows that all 23 titles he considers are still available.)

The customary flair and insight are well in evidence and the book is a pleasure to read. I have no wish to pass a long future "in the jaws of Satan", a fate pictured by Gould for several reviewers of his books, but I have two misgivings about this one. First, readers familiar with Gould's writings will feel a definite sense of *déjà vu*, for we are well aware of his limited regard (to put it mildly) for reductionism, intelligence testing, pop-sociobiology, adaptationism and so on, and we know he objects strongly to the non-reporting of null results (he and Eldredge proposed the theory of punctuated equilibrium "to

grant stasis in phylogenetic lineages the status of 'worth reporting'"). Secondly, he attacks some authors so unmercifully and relentlessly that one feels sorry that they can exercise no right of reply.

The lucidity of his arguments and the ease of his writing are as remarkable and satisfying as ever. He is like his avian namesake the jay, colourful and garrulous. Every so often comes one of those finely turned phrases (which P.G. Wodehouse also revelled in and called his "nifties"). For example, "the success of one field after a painfully slow beginning does not guarantee that all disciplines in a similar state have a rosy future; as with species and businesses, most never get very far". Or, "although stomping dinosaurs cannot make continents drift, organisms do create and shape their environment; they are not billiard balls passively buffeted about by the pool cues of natural selection". The essays are arranged in five sections, not very convincingly in some cases in spite of the lengthy justification for the groupings in the preface. The one containing sympathetic but lively sketches of "four biologists" (Barbara McClintock, E.E. Just, G. Evelyn Hutchinson and Lewis Thomas) is particularly delightful.

The title is typically Gould. Urchin is an old name for a hedgehog and his starting-point is the aphorism attributed to Archilochus that "the fox knows many things but the hedgehog knows one big thing". The way of the hedgehog is "coherence". He balls up and presents his spines to the world but can also (though only in legend) aim his darts. Gould is happy to see



David Levine

Stephen Jay Gould — colourful and garrulous.

himself in this role and to show other traits of hedgehoggy. The storm is the plethora of scientific information and the controversy and complexity it involves. As I write there is a series of letters in *The Times* noticing that some hedgehogs keep going when they sense an approaching vehicle and so escape, rather than balling up and getting squashed, and that these smarter hedgehogs are becoming more common (as a biologist would expect). I don't know which type Gould would identify with but he is quite capable of turning at the last moment to squash any runaway ideological juggernaut. □

Alan Brafield is in the Department of Biology, King's College London (KQC), University of London, Campden Hill Road, London W8 7AH, UK.



Winding down — the Hay River in Alberta, Canada, showing meanders, cutoffs, oxbow lakes, swamps and floodplain. The photograph, taken at about 6 km, is from the 3rd edition of *Modern Physical Geography* by Arthur N. Strahler and Alan H. Strahler. The new edition has increased coverage of subjects of current importance such as remote sensing and plate tectonics, as well as environmental topics including the impacts of acid rain and environmental carbon dioxide, and the destruction of low-latitude rainforests. The book is published by John Wiley, price £31.80, \$34.65; Wiley International Edition (for students) £14.95, \$11.95.

Historical portraits of the underdog

Alan Holland

Vivisection in Historical Perspective. Edited by Nicolaas A. Rupke. *Croom Helm, London/Methuen, New York: 1987. Pp. 373. £45, \$85.*

The Animal Estate: The English and Other Creatures in the Victorian Age. By Harriet Ritvo. *Harvard University Press: 1987. Pp. 347. \$25, £14.95.*

The Criminal Prosecution and Capital Punishment of Animals: The Lost History of Europe's Animal Trials. By E.P. Evans. Introduction by Nicholas Humphrey. *Faber & Faber: 1987. Pp. 336. Pbk £4.95, \$7.95.*

THE use of animals in connection with weapons testing is no new thing, it seems. Paul Elliott places the origins of experimental physiology in nineteenth-century France not in the Paris medical school, as one might expect, but in the veterinary schools of the provinces where horses, those doughty vehicles of war, doubled as experimental subjects. In this and many other ways, what William Paton calls "the erosion of simplistic ideas" is the keynote of Nicolaas Rupke's first-rate collection of essays on the history of vivisection.

The book enhances our understanding of the *anti*-vivisection movement in two ways. First, it enables international comparisons to be made: the situation in England may be plotted against that in France, Italy, Sweden, the United States, Germany and Switzerland, and we see, for example, the remarkable similarities in the social composition of the humane movement in these countries. It is unfortunate that we have no news of the Netherlands, in the light of Judith Hampson's observation that this country currently enjoys the world's most advanced system of regulation of animal experiments.

The second area of enlightenment concerns the points of contact between vivisection and matters such as the status of science, the sources of political and moral authority and, especially, the cause of women. In a particularly fine contribution, Mary Ann Elston warns against misreading this last connection. There was bound to be overlap, she argues, given that vivisection was the subject of public debate, and that the women's issue concerned precisely the place and role of women in public life. Moreover, women formed the majority in most of the big philanthropic societies, thus enacting their role as 'civilizers'; and in any case, as Elston crisply observes, "disquiet about vivisection was not confined to those who joined societies" (p.267).

Of equal importance, though, are those contributions that focus on the emergence of experimental physiology itself, and provide insight into the aims and methods of such leading figures as Magendie, Bernard, Marshall Hall and J.S.B. Sanderson. Vivisection is seen as integral to the nature and status of physiology and, later in the nineteenth century, of medicine. It was integral, too, as Rupke

observes, to the status of the practitioners themselves, those upwardly mobile scientists who caught, as it were, the fading blasts of aristocratic and religious scorn. In this way, although the nineteenth-century arguments anticipate those now current, the sources of conflict are surely somewhat different. Moreover, medical benefit seems to have been less of an incentive then than now, only coming to the fore in works such as Richard Owen's *Experimental Physiology: Its Benefits to Mankind*, published in 1882. There was also some reluctance to stress such bene-

picture of such exploitation in full spate. Her portrait of Victorian England is based on the soundest evidence — the writings of those most closely involved with animals, as found, for example, in the records of animal organizations and in the specialist journals. Her theme is the manner in which this 'animal-discourse', and the treatment of animals that it reflected, served other rhetorical functions, not least the celebration of human domination in general and that of certain social groups in particular. Specifically, she shows how popular zoology was slanted to reflect the structures of human society; how livestock and pet breeds served to celebrate the superior status of their owners; how the animal protection societies harboured anxieties about social discipline; and how menageries and game-hunting expressed Britain's imperial aspirations.

This is no mere 'reading' of the material, for her thesis is surely a response to some genuine gaps in our understanding of Victorian attitudes towards animals. Some explanation is called for of the exaggerated features bred into prize live-



Great and small — a contemporary cartoon from *Punch* entitled 'Dog fashions for 1889'.

fits in justification of vivisection. This was no doubt due in part to the absence of hard evidence but it seems also to reflect a certain sensitivity. From opposing sides, both Ludimar Hermann and Ernst Grysanowski, who come across as among the more thoughtful contributors to the debate, expressed the view that such justification was both ignoble and egoistic.

An idea long overdue for erosion is that the human exploitation of animals has its roots in Christian teaching, perhaps laced with a touch of Greek influence. The revisionist view propounded by Harriet Ritvo, and substantially borne out by all three books under review, is that an earlier view of animals which conceded to them some independence and power, and which was assailed but not overturned by Descartes, only yielded in the eighteenth and nineteenth centuries to one that justified rampant exploitation.

Ritvo paints a vivid and even exuberant

stock, which bore little relation to the needs of practising farmers. Animals could indeed be a danger to public health, but why did the *Illustrated London News* spy "dissipated canine ne'er do wells" (p.179) lurking at every street corner?

My reservation about the book is that Ritvo occasionally overplays her hand, as when she attaches significance to Darwin's use of "refuse" — just a stylistic variant, surely? — in describing the failure of wild animals to reproduce in captivity, and takes it to signal a view of the animals in question as defiant. It is hardly up to the animals whether they "breed", as opposed to "couple" (a distinction which Darwin carefully observes several times in the passage in question), and if there is a challenge here it is surely to human ingenuity rather than to human authority.

A more important point, perhaps, because it bears on a distinction between ostensible and real significance which

From *The Animal Estate*

permeates her discussion, Ritvo does not always make it clear what level of understanding of the real significance of their actions is being attributed to the agents involved. Examples are her reference to "the rhetoric of conquest that [Sir Stamford] Raffles planned to embody in the Regent's Park Zoo" (p.206), and to the desire of Londoners to "view exotic animals in chains and cages" (p.207). Londoners were anxious to view the animals; and the animals were — perforce — caged. It by no means follows that the people concerned desired to view caged animals as such. Overall, there is a hint of caricature in the picture.

What we cannot now see as anything



"Poor Pussy's scratch is as bad as her bite" — anti-rabies measures inspired this Punch cartoon of 1889.

but caricature are the 190 or so animal trials, which took place throughout Europe and North America between AD 824 and the nineteenth century. In 1906 E.P. Evans rescued records of the trials from oblivion, and we can only be grateful to Nicholas Humphrey for the timely reprint of his book.

The unlikely work of an unlikely man — an American professor of modern languages, of Welsh extraction, who lived in Germany and wrote about animals — *The Criminal Prosecution and Capital Punishment of Animals* has an endless capacity to surprise and challenge the reader. Evans records how assorted domestic animals, rodents, insects and other creatures were variously excommunicated, hanged, burned, strangled or 'knocked on the head' for crimes ranging from crop damage to homicide. What is remarkable is neither the killing of the animals nor the manner of their deaths; rather it is the fact that they were subjected to the full panoply of the law, extending, for example, to the engagement of advocates and the provision of the 'King's bread'.

Even more remarkable, in their way, are the occasional glimpses we get of mercy tempering justice. In Austria in 1519, moles were guaranteed safe-

conduct (from the predations of dogs and cats), and the pregnant females, with their young, were granted an extra 14 days leave, before being obliged to vacate their existing premises; they were, it so happened, considerably inconveniencing the local residents. In a sodomy case at Vanvres in 1750 a she-ass was adjudged to have been raped, and therefore saved from execution, after testimony as to her good character from the priest and other worthies of the town.

Humphrey suggests we see behind these trials the need to make sense of seemingly inexplicable events by redefining them as crimes, thus implying, for example, that "the pig knew very well what she was doing" (p.xxvi). Curiously enough, however, although Evans expostulates over the "irrational and absurd" nature of the trials, he too believes that "animal intelligence is capable of distinguishing between right and wrong" (p.247). In fact he holds that the capital punishment of an insane

person (and even of an ox) would be justified if it deterred other insane people (or other oxen). Moreover, it needs to be said that although Humphrey rightly directs us to the fascinating question of the thinking behind these trials (how *much* we would understand if we could understand that!), what his foreword does not prepare us for is the fact that Evans's review of the trials is but one half of a thesis. The other half is that contemporary treatment of human criminals tended towards excessive commiseration and "maudlin sympathy" (p.203) explicable only as the 'violent recoil' from previous practices — and almost equally to be denigrated. Evans was a man who recommended the sterilization of the unfit (p.221), and was someone for whom the electric chair was the last word in modern consideration for the criminal (p.210). □

Alan Holland is a Lecturer in the Department of Philosophy, University of Lancaster, Bailrigg, Lancaster LA1 4YT, UK.

Selected passages

Niles Eldredge

The Essential Darwin. Edited by Mark Ridley. Unwin Hyman, London: 1987. Pp. 271. £14.95.*

IT MUST be a source of wonder to scientists not daily engaged in evolutionary biology that the words of Charles Darwin continue to have such import in the discipline he founded. We read Darwin because he had it right, or nearly so, about many of the fundamentals of the evolutionary process. In particular, of course, Darwin was able to specify the nature of the dynamic process underlying the design of organisms — namely, "natural selection", the deterministic (if stochastic) governor of stasis and transformation of organismic phenotypes. It is instructive that he did so despite holding erroneous conceptions of the underlying processes of heredity; wisely, Ridley has included these as well as the more enduring ideas in his collection of darwinian excerpts.

Because Darwin knew that organisms vary within local populations of species — so that in a world of finite resources some would be bound to fare better than others — and because he saw that offspring tend to resemble their parents, he realized that the features of the more successful members of a generation would tend, on average, to be passed along preferentially to the next generation. The details of how the heredity process actually works are irrelevant to the basic conceptualization of natural selection. Darwin's was a

genuine discovery. Being the first on the block, so to speak, Darwin devoted his considerable talent and tenacity to re-writing the biology of his day "in the light of evolution", to borrow Dobzhansky's phrase. Much of his explorations of the implications and ramifications of evolution remain important today. Indeed, as one example, Darwin's distinction between natural and sexual selection (not articulated until the *Descent of Man*, 1871) has only recently begun to receive the attention and acceptance that it has always deserved.

For these reasons alone, Ridley's compilation of passages which, to his generally judicious eye, best spell out the important ideas in nine of Darwin's books, is most welcome. As an undergraduate, I approached the *Origin* with trepidation. Overawed by Darwin's reputation, and cowed by Victorian English usage, it took a while for the clarity of Darwin's thought and the gentleness of its expression to come through to me. I hope that, in producing this anthology, Ridley will succeed in his quest to bring Darwin closer, especially to students — whetting their appetites and inspiring them to explore Darwin further on their own. □

Niles Eldredge is at the American Museum of Natural History, Central Park West at 79th Street, New York, New York 10024, USA.

• For those who prefer the unabridged version, Pickering & Chatto, London, and New York University Press are in the process of publishing *The Works of Charles Darwin* in 29 volumes, edited by Paul H. Barrett and R.B. Freeman. Volumes 1–10 are published in 1987, and include a general introduction to Darwin's work, the zoology of the Voyage of HMS Beagle, and two essays on the foundations of the origin of the species. Volumes 11–29 are due out in 1988.

*In the United States published by W.W. Norton as *The Darwin Reader*, price \$19.95.

Slugging it out

Adrian Desmond

The Cuvier-Geoffroy Debate: French Biology in the Decades before Darwin. By Toby A. Appel. *Oxford University Press*: 1987. Pp.305. \$35, £29.50.

GOETHE was electrified by news of it; Parisian intellectuals flocked to watch it; the press eagerly reported it. Cuvier and Geoffroy were at last fighting publicly at the Académie des Sciences, after a decade of deteriorating relations and increasing disagreement over the first principles of science. The story of their clash in 1830 grew to mythic (and misshapen) proportions through constant retelling in the nineteenth century. But now Toby Appel in her long-awaited and exhilarating study has cut through this older historiography to provide the definitive modern account.

The emphasis is on 'modern', for Appel sees this as a social encounter, rather than a clash of aetherial ideas. She lays bare the scientific discord all right: the imperious Georges Cuvier with his four *embranchements* of animal life, the tetchy Geoffroy St Hilaire with his "unity of plan" stretching from monad to man; Cuvier's functional explanation of structure, Geoffroy's uniform morphological laws. Appel's revisionism is strongest in this area, and she shows that both men actually defended more extreme positions than later commentators admitted.

But she does much more, and embeds this scientific analysis in a rich study of contemporary institutional politics. Her evocation of professorial life in the Jardin des Plantes — the prima-donna personalities, shifting allegiances, rival methodologies and struggles for patronage — provides the context to make sense of the passions. Outside the Jardin, too, tensions were rising as France lurched towards revolution (it came three months later), and Appel explains why, in this anticlerical climate, Cuvier feared Geoffroy's deistic, lawful science, which "would reduce Nature to a sort of slavery" and divorce it from God's immediate action.

The Académie debate was primarily about Geoffroy's "unity of plan". (It was sparked off by his support for attempts to trace homologies between squids and fishes.) But the new "philosophical anatomy" also embraced recapitulationist embryology, the animal series, and sometimes verged dangerously close to evolution — and Cuvier abhorred all of these speculative theories for implying that nature had an "autonomous existence". Of course this explains why it appealed to the republicans, and Appel depicts Geoffroy going out of his way to canvass the popular vote. His romantic vision



"The great philosopher fell into a profound sadness" — a caricature of Geoffroy and Cuvier by J.J. Grandville. (From the book under review.)

attracted the younger writers too: George Sand revelled in her meeting with the aged Geoffroy, whom she described as "a rather curious beast, as ugly as the orangutan, as talkative as a magpie, but for all that full of genius".

It is a pity that Appel was not able to mention the wealth of work published during the past five years on philosophical anatomy in Britain (a subject which L.S. Jacyna has made his own) the more so because it supports her main conclusions: that there was no outright victory for Cuvier, that Geoffroy's supporters were more numerous than had been thought, and that comparative anatomists in the 1840s hammered out a compromise between the functional and morphological approaches. But this neglect does not detract from Appel's achievement. This newer work simply enables us to flesh out her rather *ad hoc* discussion of British biology, while supporting her main thesis.

Historians are only now coming to realize how prevalent and important morphological theories were in the pre-darwinian period. *The Cuvier-Geoffroy Debate* contributes substantially to this changing historical perspective. Toby Appel's masterly study is destined to become a landmark. □

Adrian Desmond is an Honorary Research Fellow in the Department of Biology, University College London, Gower Street, London WC1E 6BT, UK, and author of *Archetypes and Ancestors: Palaeontology in Victorian London 1850-1875* (University of Chicago Press, 1984).



NATURAL
HISTORY
MUSEUM
PUBLICATIONS

First Impressions. Early British Discovery of Australia

M. Steven

Published to coincide with the Australian Bicentenary, this lavishly illustrated book draws on the records left by artists Sydney Parkinson and Ferdinand Bauer of the difficult early days of the convict colony in New South Wales.

96pp (approx). 0 565 01023 9 £4.95

Dogs of the Last Hundred Years

K. Dennis-Bryan & J. Clutton-Brock

Documents the history of the remarkable collection of mounted dogs principally exhibited at the Zoological Museum, Tring (a branch of the BM(NH)). Each specimen not only represents its breed, but also has a personal history.

96pp, 60 b/w & 40 colour illus. Pbk. 0 565 01022 0 £9.95

A Study of Blackbirds, revised ed.

D. W. Snow

A reissue with updated bibliography and new preface of Snow's classic work on observing one species of bird over a long period of time.

192pp. Pbk. 0 565 01021 2 £9.95

Dinosaurs In The Attic. An excursion into the American Museum of Natural History

D. J. Preston

Traces the history of the collections, the characters involved and a 'behind the scenes' look at this world-famous museum.

244pp, 36 b/w photos.

Hardback. 0 312 21098 1 £14.95

(Published by St. Martin's Press, USA)

Handbooks For The Identification of British Insects

This series, published by *The Royal Entomological Society*, is now available exclusively through us.

Details available from:
Sales Department, Natural History
Museum Publications, Cromwell Road,
London SW7 5BD

An alimentary tract

John Rivers

Eat Your Heart Out: The Fallacy of the 'Healthy Diet'. By James Le Fanu. Macmillan, London: 1987. Pp. 184. £10.95.

James Le Fanu has written a book that is certain to sell, though it is more problematic whether it will succeed. For the task that he has taken on is not just that of producing another popular book on nutrition, but that of arguing against the whole direction of modern public-health nutrition — in short, of spearheading the dietary counter-reformation.

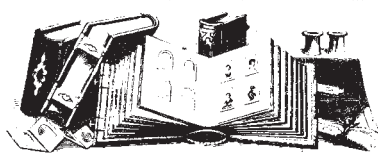
In the book he ranges against the whole modern enthusiasm for stringent eating. It is, he argues, an enthusiasm which in its dietetic Calvinism has a moral appeal that many find attractive. Moreover, in as much as its advocates have gained kudos and advancement from their stand, Le Fanu claims that there are also baser motives behind its popularity. Indeed, he argues, the only thing it has not got is a credible scientific basis.

To convince the reader of his case, Le Fanu uses a two-pronged attack. First he produces an account of the two major nutritional errors of the past — the protein fiasco which misdirected so much aid for 25 years after the Second World War, and the public-health debate of the inter-war period when the 'newer knowledge of nutrition' came to the fore and its advocates persuaded people that they had a massive problem of ill-health from consuming a poor-quality diet. Le Fanu makes the most of the ironical consequence that the public-health nutrition lobby was advocating increased consumption of the very foods it now inveighs against: butter, eggs, milk, meat and white bread.

This argument is well presented, and is all great fun. But in the last analysis it proves nothing. That nutritionists have been wrong in the past is an object lesson they should bear in mind, but it says nothing about the veracity of their current claims. Their case stands simply on the evidence adduced for it.

A critical re-examination of that evidence is the second prong of Le Fanu's attack but it is unfortunately the weaker one. Of course his central contention is right — the evidence used to back up the case for a dietary cause of degenerative disease does have many flaws. But its adequate dissection requires a more detailed analysis than he gives it. In fact, Le Fanu does not reassess the whole body of evidence used to support dietary theories of degenerative disease; rather he merely

points out the inadequacy of part of it. He almost totally ignores the evidence about polyunsaturated fat and that about the possible role of fish oils in preventing heart disease. He treats the whole lipid hypothesis as if it were nothing more than a theory linking high-fat diets to coronary heart disease (CHD). Moreover he pays little attention to the biochemical evidence, which provides plausible causal chains linking diet and disease, and concentrates on the epidemiological and the population intervention studies that have been used to investigate the aetiology and prevention of CHD. But the trouble with these data is that they are often of very poor quality. So reinterpretation cannot provide refutation of a hypothesis, though it does underline the weakness on which that hypothesis was based.



Supplementary reading

*When summer holidays are over
And youths have packed their cricket bats;
When bees can pollinate sweet clover
(Their predators controlled by cats);
When tortoises and finches vie
On evolution's apple-tree,
To flourish and diversify
Ab ovo ad posterity;
When autumn leaves begin abscission
To redden, and at length to fall,
With equinoctial expedition
Responding to Dame Nature's call.
The book reviewers start to flower,
Suffusing soft, expensive scents,
And editors in ivied tower
Compile autumnal supplements,
Selecting, for our delectation
(As far as pocket-books allow),
A groaning smorgasbord collation
Of science, our most sacred cow,
Presenting, in a final section,
Reviews of volumes — mostly new —
That, by some Natural selection,
Are rated worthy of review.
The academic bonfires burn
And old editions seek reprieves.
The avid bookworms rise and turn
The newly issued autumn leaves.
We doff our scanty summer frocks
To don more warm, unselfish jeans,
Ascending from pre-cambrian rocks,
Through fishes, reptiles, monotremes —
Through evolutionary climates,
From protoplast and platypus,
Through curates and canonic primates —
To evolution's summit: Us.
And now, with minds and bodies bent,
We welcome, in our new position,
This supplemental monument,
To scientific erudition.*

Ralph Lewin

Nutritional epidemiology has a weak point at the level of dietary assessment, which Le Fanu misses. Often the food intakes which are cited are merely apparent food consumption figures grossly aggregated for groups, sometimes even nations. Even real dietary intake data can be challenged as they tend to be of dubious accuracy, often coming from studies of small groups for at most a few weeks. Such data make poor predictors of long-term dietary patterns, and it is these which are supposed to be the cause of degenerative disease. So just as these data are suspect when used in support of the conventional wisdom, they are suspect when used to attack it.

The population interventions that have not worked cannot be interpreted either. They show that such interventions have had only limited success, but not whether the failure reflects a failure of the population to stick to the dietary advice, or whether that dietary advice was no good.

Some evidence does present problems that Le Fanu ascribes to it — for example, the fact that after CHD mortality had risen inexorably from the beginning of the century it then began to decline from about the early 1960s in countries as diverse as the United States, Australia and New Zealand, and the fact that this trend has spread eastward since then so that recently even the British statistics have begun to decline (though Eastern Europe has yet to be affected). This temporal change is, as Le Fanu argues, difficult to square with any dietary hypothesis except by special pleading of a purely unprovable kind. But it is not true when he says that it has occurred despite the fact that dietary fat intakes have remained constant; in the United States at least, energy intakes have tended to decrease during that period so that, on the same fat content of the diet, the fat intakes have declined.

Le Fanu does not help his case by taking a cavalier attitude towards the data he re-analyses. For example, without comment he moves from citing mortalities as gross figures to rates for specific age-sex groups; and he routinely fails to attribute the sources of his data, and in some graphs dispenses completely with scales on the axes. Le Fanu is not writing for professionals, but it is hardly appropriate that an author who accuses scientists of bamboozling the public should use data in a potentially misleading way.

Overall this is a book that may fuel the debate about diet and disease, but it will not add substance to the legitimate doubts about current orthodoxy. Nevertheless if it stimulates a critical reassessment of that orthodoxy, it will have served a purpose. □

John Rivers is a Senior Lecturer in the Department of Human Nutrition, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK.

Hopes and fears and a national phobia

Daniel S. Greenberg

The Dread Disease: Cancer and Modern American Culture. By James T. Patterson. Harvard University Press: 1987. Pp.380. \$25.95, £20.75

In 1970, the US Congress unanimously resolved that cancer should be cured by 1976 "as an appropriate commemoration of the two-hundredth anniversary of the independence of our country".

Hallucinogenic declarations are common on Capitol Hill. Little noticed, the decree stood among many posturing diversions from serious legislative business, a Democratic-inspired jab at Nixon's penny-pinching White House. But this *was* serious business. The subject was the most feared disease in the nation. The goal was the defeat of death, a universal favourite, but one especially cherished in yogurt-gulping, jogging American culture. Adoption of the resolution proved to be a major step towards passage of the National Cancer Act of 1971, which was quickly embraced by a previously opposed Nixon when it was bound to pass. (He attributed his interest to the death of an aunt by cancer.) Declaration of the "war on cancer" was followed by extraordinary sums of government research funds, some \$15 billion so far, making cancer the most richly financed field of research outside the blank-cheque realms of defence and space.

Good for cancer research — better there than in the aforementioned, many have said. But the abundance of money is difficult to justify in terms of the health needs of the American people. The death rate from cardiovascular diseases is double that from cancer, which remains predominantly a disease of the elderly. At the same time, the United States lags behind a dozen or so nations in controlling infant mortality. And some 37 million Americans lack insurance for health care, very little of which is given away in this land of free enterprise for the poor and subsidies for the rich.

Furthermore, the cancer establishment, until recently, has concentrated its wealth on a desperate quest for cures, while neglecting preventive measures. Backed by the US Treasury, the anything-goes, curative strategists have, for example, scooped up hundreds of thousands of samples of mud and plant matter from around the world in a futile search for nature's own cure for cancer. Meanwhile, the cancer warriors were late arrivals to the war on smoking.

How did cancer, among humanity's many afflictions, come to be so feared in the American psyche? And how did a self-perpetuating cancer establishment become so privileged, powerful and selective in the dispensation of immense

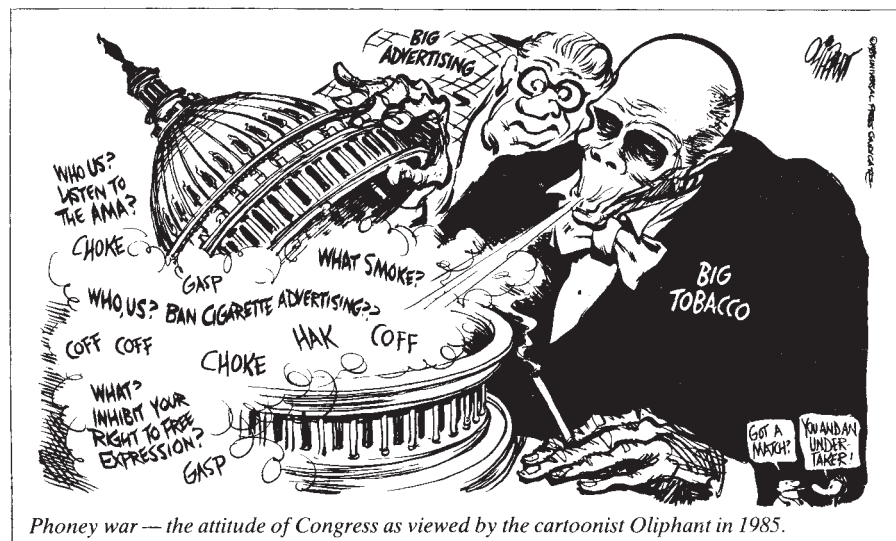
biomedical resources?

These questions are addressed by Professor Patterson, a historian at Brown University, in *The Dread Disease*. This is an ambitious and original attempt to advance past the bountiful literature on cancer politics produced in recent years by scholars, journalists and public-health activists. Patterson is only peripherally concerned with the biology and treatment of cancer and even less so with the clockwork of cancer legislation, for which Richard A. Rettig's *Cancer Crusade: The Story of the National Cancer Act of*

cancer — a declaration of "war". Starting in the late nineteenth century, many elements started to come together, including growing expectations about the power of medicine and an uncritical press that relished tales of heroic scientific advance, and still does. (Patterson's collection of 'cure' reports from the serious wing of the American press illuminates a disgraceful chapter in journalism.)

But in the elevation of cancer into a national phobia, perhaps most important of all, Patterson states, was that "The denial of death was especially strong in the United States, land of perceived opportunity, technological progress, and economic growth" (p.32). Especially after the high-tech triumphs of the Second World War, Americans — now, for the first time, comfortable with government as problem solver — couldn't tolerate cancer.

In some of the most perceptive writing



Phoney war — the attitude of Congress as viewed by the cartoonist Oliphant in 1985.

1971 (Princeton University Press, 1977), remains the standard work.

"Rather", Patterson explains,

I concentrate on elucidating the ways that the dread disease has reflected social and personal concerns during the modern, industrial era of United States history: concerns about illness, health, medical practices, and death and dying . . . The history of cancer in America is also a history of social and cultural tensions . . . Many [Americans] clung defiantly to ideas about cancer that were scorned by the majority of scientists. Still others were simply terrified: the persistence and growth of popular cancer-phobia are central themes of this book [pp. viii, ix].

Why cancer as the most feared disease, the darling of federal health research, the most popular object of medical charity, and the specific interest of very first institute to be founded in what eventually evolved into the great National Institutes of Health? The obvious part of the answer is that cancer has earned its horrifying reputation. But, as Patterson skillfully demonstrates, more than fear was needed to inspire America's unique reaction to

yet produced on America's response to cancer, Patterson describes in detail the painful dichotomy between the victorious communiqués that routinely emanate from the cancer establishment and common personal experience with the disease. "In challenging the optimists", he concludes, "disillusioned Americans did not need to look very far for evidence".

Prevention is now the central theme of America's war on cancer, with the National Cancer Institute dedicated to halving the death rate by the year 2000. Dietary modification and cessation of smoking are among the key elements in this strategy. Both, of course, could have been instituted as national policy 20 years ago. Why they weren't, and why the cancer-terrorized masses in the United States long tolerated the warped strategy against cancer, are well explained in this fine, important book. □

Daniel S. Greenberg, 3736 Kanawha Street NW, Washington, DC 20015, USA, is editor and publisher of the newsletter Science & Government Report, and a long-standing commentator on the "war on cancer".

Seeing the wood for the trees

Norman Myers

The Purpose of Forests: Follies of Development. By J. Westoby. Basil Blackwell: 1987. Pp. 343. £35, \$60.

TROPICAL forests are falling; we all know that. Yet it is simplistic to blame the feller of the forests, whether commercial logger or subsistence peasant. Both act in response to economic and social forces which they may only dimly discern, let alone understand. So to point an accusing finger at the man with the chainsaw or machete is to aim at the symptoms of the problems, rather than the problems themselves.

But today the problems have grown so pervasive and profound that it may even be of little use to tackle them any more. Rather we should address their sources. Within most tropical-forest countries, these sources include development strategies gone wrong, economic inequalities, oligarchical governments, absurdly inefficient administrations and the Cinderella status of forestry departments.

In the Philippines, until just a couple of years ago, more than half of all logs



Holt Studios

Visible symptoms — deforestation of the South Thailand hillsides.

exported were "poached" by cronies of former president Marcos. Similar timber rustling is common in Thailand. In Madagascar, the parlous state of the economy has been further aggravated by debt burdens, closing off prospects of making agriculture more intensive in traditional farmlands — this being the soundest way to reduce the "push" motivation for impoverished peasants to migrate towards the remaining forests. In Brazil, the best strategy to safeguard Amazonia is not to build a fence around it in order to keep out throngs of landless settlers. It is to engage in agrarian reform in southern Brazil, with its major agricultural lands, a region which could accommodate twice as many farmers as there are now in the whole country.

These are the sorts of insights proclaimed by Jack Westoby. The 16 lectures and papers that make up this book reflect the evolution of his views since the early 1960s, when he was rising towards the upper echelons of the UN's Food and Agriculture Organization's Forestry Department. In those far-back days it was supposed that tropical forests were little more than so many board-feet of timber; the approach was "How can we best exploit the forests?". By the 1970s a new view became current, to the effect that the forests contribute not only material goods but environmental services as well, notably in their role as watersheds. So the approach changed to "How can we best develop forests?", even if that sometimes meant leaving them standing. By the 1980s there was a further shift to "How can we enable forests to make their full contribution to development overall?" — that is, how can they best add to a nation's gross national product?

Yet even this last approach, enlightened as it is when compared with earlier times, misses the point. As Westoby demonstrates in a dozen graphic ways, forestry is not primarily about trees, it is about

people. It concerns trees only insofar as they serve the legitimate needs of people — especially the multitudes of half-destitute people who share the rural habitats of forests. If these people do not perceive forests as sustaining their welfare, they will take a matchbox to them. Unless this central premise is accepted, there will soon be no tropical forests left for foresters — or conservationists — to agonize over.

Telling professional foresters that they should be concerned about local villagers as well as trees has not always endeared Westoby to his more traditionalist colleagues. When he presented a keynote address at the World Forestry Congress in Jakarta, and took to task the assembled thousands for misperceiving their mission, he caused many an eyebrow to be raised in askance. But it is one of the strengths of this book that Westoby is always briskly straightforward. Thus we find the substance is set out with fierce conviction. And he always presents his acerbic message with genial goodwill: he takes his subject seriously, himself much less so.

By the time I joined FAO's Forestry Department, Westoby had retired, and for all I knew his message had been quietly filed away in the archives. So I especially welcome this book, and I am sure I shall return to it many a time for inspiration as well as instruction. There is little in it about cutting rotations, species provenances, wood chemistry and other technical matters. But there is much in it to commend to the development expert, resource economist, land-use planner, academic, conservationist — and the forester who recognizes the radically changed nature of the world in which he pursues his profession. □

Norman Myers, Upper Meadow, Old Road, Headington, Oxford OX3 8SZ, UK, is a consultant on development and environmental matters.

THE MANAGEMENT OF AIDS PATIENTS

Edited by DAVID MILLER, JONATHAN WEBER and JOHN GREEN

The Management of AIDS Patients is the first comprehensive guide to the practical clinical management of patients with AIDS or HTLV III infections. The book avoids the sensational aspects of the disease, offering solid advice and information for all people involved in patient care.

The editors and many of the contributors come from St Mary's Hospital, London one of the foremost centres in Britain for the treatment of AIDS patients. The knowledge and experience of these experts has been combined to provide a much-needed book for doctors, nurses, dentists and indeed for health-care professionals everywhere.

1986

200pp Hardback 0 333 40465 3 £30.00
Paperback 0 333 40466 1 £12.95

Please order this title from your bookseller — or in case of difficulty from Felicity Davie, Macmillan Press, Houndmills, Basingstoke, Hants RG21 2XS.

M MACMILLAN PRESS

Sea versus PC

William H. Press

A View of the Sea: A Discussion Between a Chief Engineer and an Oceanographer about the Machinery of the Ocean Circulation. By Henry Stommel. Princeton University Press: 1987. Pp. 165. \$19.95, £12.60.

HENRY Stommel, who is now approaching the age of retirement, is an extraordinary figure in post-war physical oceanography. He was responsible for the foundations of our modern understanding of oceanic circulation, and is famous as a raconteur, tinkerer, amateur explosives maker, backyard railroader. More importantly, he is celebrated for an uncanny physical intuition about how the ocean works.

A few years ago, Stommel began to write fascinating, off-beat, popular books: one, *Volcano Weather* on the so-called Year Without A Summer (1816), another on islands which appear on historical nautical charts, but subsequently vanish. No surprise, then, that the Sloan Foundation should ask him to contribute a personal autobiography to their distinguished series. In these times, scientists no less than starlets are encouraged to bare all.

Stommel did decide to write a personal book, but not under the Sloan Foundation's auspices. *A View of the Sea* is personal, after a fashion: cranky, idiosyncratic, uncompromising. There is little more than an occasional paragraph or two of autobiography in this book, and only a tease of shipboard anecdote. Stommel's idea of a personal book is not to take us into his life, but rather to try to take us into his mental spaces, to show us exactly *how he thinks* about ocean circulation. The result is fascinating, even when it can hardly be termed accessible.

Stommel thinks not in equations (certainly not in partial differential equations), but rather in pictures and force diagrams. Slabs of ocean water overlay other slabs. They tilt, slide on each other, thicken, thin, react to the strong constraints of the Coriolis force and the geometry of the two-sphere. "Then, by Rule 2, I found the GV [geostrophic velocity] in a relative to *b*. Now, Rule 3 demands that the actual GV in layer *a* (let's call it GV_a) must be parallel to the PT [potential thickness: thickness divided by sine of latitude] lines in *a*, and that GV_b must lie parallel to the PT lines in *b*." There are pages and pages of this kind of thing.

What does one learn by working (I do mean *working*) it through? Quite a lot, if, like me, you are not an oceanographer. Because of the Coriolis force, ocean water does not flow in the direction that the wind pushes it, but at right angles to that

direction. Westerlies in the North Atlantic (which blow to the east) thus drive water south; but trade winds to the south blow to the west and drive water north. The water has nowhere to go but down (this is Ekman pumping). Then where does it go? To accept the further fluid pumped down on top of it, and to be geostrophic, it has to move to a region where the Coriolis force is weaker, that is, south.

This cold, broad layer of North Atlantic water continues to be driven south. It goes deeper, and spreads laterally, as successively warmer layers are laid down on top of it. Conservation of mass must require a reverse, northwards, flow somewhere. However, geostrophic flow and conservation of potential vorticity (Stommel has his own odd way of explaining this) explicitly forbid such a return. Something has to give, and it does. At the western edge of the basin there can be a dissipative boundary layer which breaks some conservation laws. The result is the Gulf Stream (in the Pacific, the Kuroshio), with a fast core jet that is able to cross isobars (violate geostrophy), and transport vorticity internally. It all fits together. Stommel calls it a Chinese puzzle.

In Chapter 11, the narrative takes a

bizarre and unexpected turn. The author announces that on Thanksgiving Day 1984 he bought himself an IBM/PC-compatible personal computer, with a colour/graphics card and spinwriter. He then devotes the rest of the book to the details of nine computer programs, printed in BASIC in the back of the book, for calculating ocean flows. The programming style is unstructured and virtually unreadable. We are caught in a melodrama of grand proportions: just when (at no small investment of concentration) we have come to appreciate, and have begun to understand, our narrator's deep geometric insight, he becomes infected with a classic, perhaps terminal, case of PC addiction. Insight is thrown away. He is transformed into, simply, a hacker.

There is thus a kind of moral lesson in this book. A diskette with Stommel's programs is available separately. The programs do run. One of them is babbling away on my computer screen right now. □

William H. Press is Professor of Astronomy and Physics in the Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, Massachusetts 02138, USA, and co-author of *Numerical Recipes: The Art of Scientific Computing* (Cambridge University Press, 1986).

The ice man

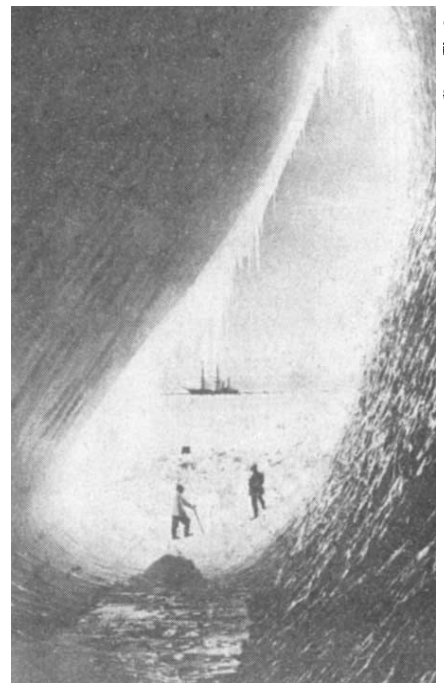
Derek Fordham

The Ice: A Journey to Antarctica. By Stephen J. Pyne. University of Iowa Press/Arlington Books, 15-17 King Street, St James's, London SW1: 1986-1987. Pp.428. \$37.50, £12.95.

A NACREOUS mass spins in the gyre of the Antarctic Ocean; almost imperceptibly it hesitates and disappears in a veil of fog and pack ice. This is ice, the ice of the title and *The Ice* (always in capitals) which is the book's leitmotiv. All its forms, from frazil to vuggy, are encountered and meticulously described as succeeding pages follow the gyrating iceberg from diamond dust to disintegration.

At first, mankind has little place in Stephen Pyne's narrative; ice, we are told, is the only component of Antarctic geography. Certainly the author's story owes little allegiance to Apsley Cherry Garrard's philosophy, expressed in *The Worst Journey in the World* — perhaps the best book ever written on Antarctica — that it is the spirit of man that has shaped that continent.

Man is conceded one resounding achievement — his success in establishing a tenuous foothold on the continent. Pyne's intense awareness of the special nature of man's relationship with *The Ice* has distinct parallels with the views of the



In perspective — the Antarctic exploration ship Terra Nova seen from an ice cave. Photograph taken by Herbert Ponting.

geographer, Yi Fu Tuan, who wrote that, "in open space once can become intensely aware of a place, and in the solitude of a sheltered place the vastness of space acquires a haunting presence". This presence has inspired Pyne to write intense and perceptive sections on the oceanography, geology, glaciology, art,

From *The Ice*

literature and geopolitics of the Antarctic in prose that is at times powerful but also, over 400 pages, can occasionally be fatiguing.

It is The Ice, the forces which create and destroy it and the forces which it in turn influences, that dominate the book. Because over 60 per cent of the world's fresh water is locked into the Antarctic, and tabular bergs the size of Belgium are not uncommon, perhaps this preoccupation is understandable, if not particularly stimulating for the reader.

It is only where man makes an appearance, and the ice terranes and ice orrerys are relegated temporarily to the background, that a wider perspective becomes apparent. Pyne develops the thesis that The Ice imposed on explorers, as on artists, the need to develop not only special technologies but also to adopt new casts of thought. These concepts are analysed and expanded at some length in the book, but the result seems to offer little of substance

to the scientist, and the humanistic perception, with which the author is credited on the jacket, although evident may well confuse the layman.

The author claims to have found in the Antarctic ice, "from core to margin, from polar plateau to open sea an allegory of matter". His thoughts about his allegorical discovery in, and return from, the icy vastness of the Antarctic seem to be encapsulated in the lines he quotes from Edgar Allen Poe's "Dreamland":

I have reached these lands but newly
From an ultimate dim Thule
From a wild weird clime that lieth sublime
Out of SPACE and out of TIME.

The space and time of the hauntingly beautiful Antarctic are there to be glimpsed in this, the author's own dreamland. □

Derek Fordham, 66 Ashburnham Grove, Greenwich, London SE10 8UJ, UK, is Secretary of the Arctic Club and has travelled widely in polar regions.

When dark is light enough

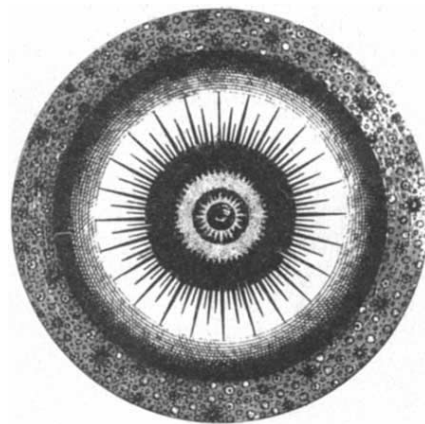
Owen Gingerich

Darkness at Night: A Riddle of the Universe. By Edward Harrison. *Harvard University Press: 1987. Pp. 293. \$25, £19.95.*

"A GREAT cause of the night is the lack of the Sun", says the innocent shepherd in *As You Like It*, and though we may chuckle at his naiveté, few of us stop to ponder the more profound alternative. "A great cause of the darkness at night is the lack of Suns". If the Universe is infinite and randomly filled with stars, then any line of sight must eventually reach a star, and the vault of the heavens should be aglow with the brilliance of an average star's surface.

If you are convinced that the Universe is infinite and randomly filled with brightly shining stars (or galaxies of stars), then the darkness is a paradox. But if you believe that the Universe is either limited or not randomly filled with stars, then the darkness is a riddle to be solved. Hermann Bondi, who in the early 1950s accepted an infinite steady-state Universe, popularized the problem as "Olbers' paradox". He claimed that the red shift of the distant galaxies solved the puzzle, and that earlier astronomers could have deduced the idea of the expanding Universe merely from the darkness of the sky at night.

Since Bondi's *Cosmology*, "Olbers' paradox" has been part of the astronomers' vocabulary, and the problem of darkness at night has been the subject of



Wheels within wheels — Thomas Wright of Durham suggested in his Original Theory or New Hypothesis of the Universe (1750) that the stars occupy a shell around a mysterious galactic centre. (From the book under review.)

several books, though none as beguiling as this new offering from Edward Harrison, the cosmologist and perceptive historian of science from the University of Massachusetts.

Harrison's account is beguiling not because it lures the reader into believing falsehoods — indeed, it looks as if he has finally solved this antique problem correctly — but because it is done so cleverly that the reader feels he is really understanding the solutions with only the barest hint of mathematics. Harrison runs a veritable Cook's tour through the history of astronomy, always with the unique perspective of the darkness riddle. His tour bus rushes on, with occasional diversions, from antiquity to Einstein, past Kepler, Kant and Kelvin, Cheseaux, Charlier and Fournier d'Albe — with stops at 15 possible solutions to the mystery.

The prize goes — somewhat ambiguously — to the remarkable insight of Edgar Allen Poe and to the independent and nearly forgotten calculations by Lord Kelvin. Harrison has resurrected both of these accounts, and he accepts their answer: stars simply do not shine long enough for there to be the necessary infinite number to dazzle the night-time sky. He describes how the Bondi–Gold red-shift solution would indeed work for a steady-state cosmology but not for the big-bang model. At the last stop on the tour, Harrison recasts the Poe–Kelvin solution into a somewhat different form in order to show that there is an even better way to look at the problem. There simply is not sufficient energy in the Universe for the stars at night to be so many and bright.

As a cosmologist, Harrison seems to have withstood the critical examination of his peers, for his solution has appeared without refutation in technical journals. The extensive bibliography as well as the success of his detective work attests to his skills as a historian of science. But the swift, nutshell descriptions of astronomers of yore will give chills to sterner, more professionalized historians. I found the chaotic inconsistency of book titles, sometimes in their original tongues and sometimes in translation, rather careless, and I was distressed that an appendix devoted to "Digges on the Infinity of the Universe" nowhere states that this is, except for one small but crucial part, merely Digges's translation from Copernicus. And to label as "Stoic" any astronomer who accepted a bounded system of stars within unlimited empty space — such as Harlow Shapley in 1920 — is less than helpful.

Harrison is quite forgiving that Bondi popularized the "riddle" under the misnomer "paradox", and that he assigned it to Olbers instead of Cheseaux, who had specified the problem earlier. Indeed, Olbers had Cheseaux's book in his library, but must have forgotten about the Swiss astronomer's account of 1744 when he wrote his own paper in 1823. Referring to Bondi, Harrison says, "Encumbered with one or two trifling historical inaccuracies, the riddle gained a new lease on life". His relaxed attitude may now stand him in good stead: since this book went to press, Frank Tipler has shown that Poe's solution of the riddle was given even earlier by the selenographer Johann Heinrich Mädler, and even that Poe knew of Mädler's speculations on this subject. Such are the pitfalls in an attempt at historical completeness! Nevertheless, Harrison's tour is nearly a *tour de force*, and certainly a fine example of science popularization. □

Owen Gingerich is a Professor in the Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, Massachusetts 02138, USA.

A question of great impact

John F. Kerridge

Coon Mountain Controversies: Meteor Crater and the Development of Impact Theory.

By William Graves Hoyt. University of Arizona Press: 1987. Pp.442. \$40, £34.50.*
Meteorite Craters. By Kathleen Mark. University of Arizona Press: 1987. Pp.288.
 \$29.95, £24.50.*

TRACING the evolution of meteorite impact theory affords us several insights into why scientific progress can prove at times to be so painfully slow. Thus, we can observe dogma characteristically defying logic: "[Impact theory] is based on assumptions which have no support either from what we see going on around us or from anything which history or scientific investigation tells us has happened in the past". The boundaries erected between disciplines are seen interfering with the flow of information: "But why there are not more [impact craters on the Earth], or at least some evidence of their remains

extent. Although Mark aims for a global, indeed Solar-System-wide, and non-technical overview of the emergence, growth and consolidation of impact theory, the focus of Hoyt's book, as its title implies, is on the wealth of scientific, historical and even financial detail concerning Meteor Crater, formerly known as Coon Mountain, and its indefatigable explorer from 1903 to 1929, Daniel Moreau Barringer.

The plural in Hoyt's title is important. For him, as for Barringer, the struggle to prove the impact origin of Meteor Crater was only part of the story; the attempt to

the vertical impact trajectories apparently required. For Barringer, a lifetime's familiarity with high-powered hunting rifles provided an adequate analogy: the "splashing distribution of ejected material" from a shot into "stiffish mud" led to a round crater, even at an incident angle of 45°.

That analogy sufficed for Barringer during the first controversy but it let him down badly when the debate turned to the mass of projectile remaining beneath Meteor Crater, a quantity of vital importance to his future mining fortunes. It turns out that the physics involved in the excavation of hypervelocity-impact craters is far too complex to be handled by simple argument-by-analogy. Lacking, in common with virtually all other scientists of the time, an understanding of that physics, and resistant to the idea of a process that would have vaporized or disintegrated his precious ore, Barringer failed to recognize the literally explosive significance of circularity so that he grossly overestimated the mass of residual meteoric material. The bullet had, in fact, vanished from the target.

Hoyt blends the scientific issues, the commercial and legal factors, and the personalities involved into a sure-footed narrative that never fails to hold the reader's interest. As a chapter in the history of science, it is of considerable,



From Coon Mountain Controversies

A hell of a big hole — Meteor Crater in Arizona, over 1 km in diameter.

... has not been explained", according to an astronomer in 1919. We are reminded once again of the perils of failing to quantify the quantifiable: "It's a hell of a big hole and it was made by a hell of a big thing", compared with "... the mass [of the Canyon Diablo meteorite] cannot have exceeded 3,000,000 tons and ... may well have been of the order of 100,000 tons". And we can feel across the years the force of a confident assertion: "It is certain that no explosive theory could produce the terraces such as we see in [the lunar crater] Copernicus".

Those quotations can be found in *Coon Mountain Controversies*, the more academically orientated of two recent books on the development of impact theory, both emanating, remarkably enough, from the same publisher. In fact, although there is considerable overlap of material between the books, they differ so markedly in scope, depth of detail and level of presentation, that their readerships are unlikely to overlap to the same

prove its value as a mineral resource was an even more dramatic and, in Barringer's case tragic, affair. In practice, the two controversies (impact versus endogenic origin, and large, slow bolide versus small, fast one) had rather little in common. However, both hinged, to a certain extent, on the interpretation of a well-known feature of impact craters, most noticeable on the Moon, namely their circularity. It is instructive to examine Barringer's approach to this issue.

Until about the time of the First World War, for most scientists, geologists and astronomers alike, the circularity of virtually all lunar craters represented one of the main obstacles to acceptance of an impact theory for their origin. Most meteoroid impacts on the lunar surface being oblique, it was believed that the resulting craters would have been predominately elliptical. Even G.K. Gilbert, a figure whose almost mythical shadow falls heavily, though enigmatically, across the pages of Hoyt's book, felt it necessary to invoke a cockamammy swarm of providentially located "moonlets" to supply

Copies of articles from this publication are now available from the UMI Article Clearinghouse.

Yes! I would like to know more about UMI Article Clearinghouse. I am interested in electronic ordering through the following system(s):

- ☐ DIALOG/Dialorder ☐ ITT Dialcom
☐ OnTyme ☐ OCLC ILL Subsystem
☐ Other (please specify) _____
☐ I am interested in sending my order by mail.
☐ Please send me your current catalog and user instructions for the system(s) I checked above.

Name _____
 Title _____
 Institution/Company _____
 Department _____
 Address _____
 City _____ State _____ Zip _____
 Phone (____) _____

UMI Article Clearinghouse

Mail to: University Microfilms International
 300 North Zeeb Road, Box 91 Ann Arbor, MI 48106

*In Britain available from Eurospan, 3 Henrietta Street, London WC2E 8LU.

and probably lasting, value; it is difficult to imagine a more carefully documented and sensibly reasoned account of the way in which ideas on impact theory evolved during the era covered by the book. If it has a failing it is that there is little attempt to put the issue of the acceptance of impact theory into a broader philosophical context. For example, was the opposition on the part of traditional geologists inherited from Lyell, or was it more a reaction to ideas arising outside the mainstream of science?

Compared with Hoyt, Kathleen Mark paints on a much broader canvas, paying more attention to the controversy surrounding "cryptovolcanic" structures, to the increasingly sophisticated study of the erosional pathology of impact craters, and to the continuing search for physical criteria for an impact origin and the structures with which they are associated. To render such material at an avowedly non-technical level is an ambitious undertaking but in this case it has been

successful. The author handles the basic scientific concepts with confidence and describes them with clarity. Many students, nominally more advanced than the audience for which the book is designed, will nonetheless find it instructive.

Both books do justice, in their different ways, to the field of impact theory. Both authors handle observations and ideas with aplomb, but it's probably fair to say that their accounts also tend to neglect the contributions of experimental science. For example, the hypervelocity-impact experiments so beautifully designed and executed by Don Gault surely deserved more than just a footnote mentioning their reproduction of shatter cones. Such carping aside, however, these are excellent books; one or other of them should meet the needs of almost anyone wishing to learn, or learn more, about the history of impact theory. □

John F. Kerridge is at the Institute of Geophysics and Planetary Physics, University of California, Los Angeles, California 90024, USA.

Light industry

John N. Howard

Land's Polaroid: A Company and the Man Who Invented It. By Peter C. Wensberg. Houghton Mifflin: 1987. Pp.258. \$18.95.

EDWIN Land was born in Connecticut in 1909. Even as a youth he had decided he would achieve fame and fortune in the world of science and technology. After all, had not Thomas Edison, Henry Ford, Alexander Graham Bell and George Eastman — all self-made men, and all then still alive — succeeded by developing practical, usable technology?

When he was 17 his parents sent him to Harvard to study physics. Before the first year was over he had already picked a likely area for a scientific breakthrough: polarized light. Unfortunately he could not immediately see any important practical application. Then, on a vacation trip to New York City he walked into Times Square, saw the bright lights and the glare of automobile headlights, and recognized the need, the opportunity — suppose automobiles had headlights emitting light in one plane of polarization, and windshields of the crossed plane; then, forward vision at night should be unimpaired, but dangerous glare from oncoming headlights would be filtered out.

Land dropped out of college, took a room in New York, and spent every day at the library reading every published paper about polarized light. An earlier researcher, William Herapath, had found that tiny crystals of iodoquinine sulphate strongly polarized light, but Herapath had been unable to grow them large enough to be useful. Land repeated some of these experiments and determined that he could make the needle-like crystals line up in a magnetic field. Perhaps he could then press them into a sheet of film.

He went back to Harvard to better laboratory facilities, but just short of graduation dropped out again, this time

(with a physics colleague) forming a small company to make polarizing film, which they named Polaroid. By 1935 he had orders from Eastman Kodak for polarizing filters for cameras. Then the American Optical Company introduced Polaroid sunglasses. The business grew, incorporated in 1937 as the Polaroid Corporation, and during the war engaged in defence research and became a successful small company. But the big money of polarized automobile headlights did not materialize: the industry feared that filters on headlights would make them dim.

For Land, the big breakthrough occurred in late 1943, while he was vacationing with his family in New Mexico. He took some pictures of his three-year-old daughter Jennifer and she wanted to see them right away. "They're not ready", he said. She replied, "Why can't I see them now?". Land went for a walk, and within an hour he had in mind the entire process of what needed to be done to make self-developing photographic film possible. The new process of "instant photography", with a working model of the camera, was demonstrated at the February 1947 meeting of the Optical Society of America, and the camera became an instant commercial success.

The company grew and prospered; in 1963 colour film for Land cameras was introduced and in 1972 Polaroid introduced the very sophisticated SX-70 camera, with automatic focusing, automatic aperture for correct exposure and a self-developing print. Through all of the expansion of the company — from a half-dozen employees to as many as 20,000 — Land remained completely in charge: chairman, president, chief executive officer and chief scientist. Starting in 1975, he gradually relinquished these posts, and retired from Polaroid in 1982 at age 73.

Land is an inventor; his name appears on 533 US patents, second only to Thomas Edison's 1,093. But whereas there are 23 biographies of Edison in print, Wensberg's book is only the second sketch of Land. The book is unauthorized, but it is not by an outsider; Wensberg worked at Polaroid for 24 years, the last several of them as executive vice president. When he told Land he was writing a book about Land and Polaroid, Land said, "That's very disturbing, Peter. Why would anyone want to read such a book?". Wensberg replied, "There's a tremendous interest now in entrepreneurship". "Isn't that an awful word", was the response, "It reminds one of a man standing behind a



Edwin Land, diffident extrovert, demonstrating "instant photography" in 1947.

Associated Press

pushcart of bananas".

Wensberg portrays Land as two personalities: as an extrovert he presided at the annual meetings of Polaroid stockholders, which were extravaganzas touting his latest (and invariably best) invention, complete with hyperbole and hoopla; but on the personal level, he is diffident and retiring — even reclusive — at least to outsiders.

In writing the book, Wensberg pored over letters and speeches by Land and material from Polaroid annual reports, and collected anecdotes and reminiscences from Land's colleagues and friends. Wensberg narrates key events in Land's life — using so-called "new journalism" techniques — by reconstructing

Land's thoughts and words in conversation. "There are some of Land's words in those stories", he says; "I confess to having enlarged upon them". The prospect of reading reconstructed quotations attributed to him drives Land up the wall: "That's a terrible indignity", he is reported to have said, "I always thought quotation marks were a sacred signal of meticulous honesty".

Honest or not, these "corroborative details add artistic verisimilitude" to a fascinating account of Edwin Land's remarkable achievements. □

John N. Howard, 7 Norman Road, Newton Highlands, Massachusetts 02161, USA, is Editor Emeritus of Applied Optics, and also Editor of Optics News, both journals of the Optical Society of America.

Logical conclusions

Philip Kitcher

Quiddities: An Intermittently Philosophical Dictionary. By W.V. Quine. *Harvard University Press: 1987. Pp. 249. \$20, £15.95.*

OVER half a century ago, a young Harvard professor of logic and philosophy published an article expressing doubts about some of the central themes of logical positivism, then in the flush of its new dominance of the Anglo-Saxon philosophical world. That article, "Truth By Convention", was to begin a series of explorations and critiques that would decisively alter philosophical views about logic, mathematics, language and science. Its author, W.V. Quine, would become the dominant figure in post-war English-language philosophy, first in the United States, and later in England, Scandinavia and Australia. Just as Quine is quoted on the back of the abridgement of the classic *Principia Mathematica* of Whitehead and Russell — "This is the book that has meant the most to me" — so generations of philosophical logicians, philosophers of language and philosophers can testify that some essay, collection or book of Quine's has been the decisive influence on their own professional researches.

Quiddities is the work of an author who has faith in his own idiosyncratic enthusiasms. Ranging from lucid expositions of philosophical topics that are central to the fields that have intrigued him throughout his career — particularly logic and the philosophy of mathematics — to entertaining and informative details about the quirks of natural languages, it fuses wit with instruction. Fundamental aspects of twentieth-century logic and set theory are explained in ways that will be comprehensible to beginners (while whetting their appetites for further details) and that will be satisfying to the initiated. Quine's



W.V. Quine — idiosyncratic enthusiasms.

talents for exposition are amply displayed in articles on familiar Quinean topics: classes versus properties, classes versus sets, definition, Gödel's theorem, identity, impredicativity, infinite numbers, natural numbers, paradoxes, predicate logic, reference and reification, singular terms, truth, universals, use versus mention, variables. For those who are unaware of the achievements of modern logic, this sequence of articles offers a beautiful presentation of the main staples of contemporary logic, as attractive as it is nourishing.

Perhaps the chief delight of *Quiddities* lies in the tidbits that are sprinkled throughout the menu. Quine has always been entertained by ordinary language, by the odd relationships among words in different languages, by the ways in which languages change, and the scraps and vestiges that are left behind. *Quiddities* reveals some of the learning that a lifetime's fascination with words can produce. Readers may be surprised to learn how the violoncello got its name, intrigued by the possibility of using inflections to build ever more complex Latin verbs, struck by Quine's demonstrations that gender in language can sometimes be used to disambiguate and can sometimes lead to intra-linguistic tension, and frankly repentant about their past usages

of 'agenda' and 'media' (both are Latin plurals corrupted as singulars). Quine wears his learning lightly, and the grace of his discussions reflects his love of the whimsy of language.

Quiddities is an "intermittently philosophical dictionary" in two senses. Not only is the philosophy punctuated by articles on matters linguistic, but the treatment of philosophical issues is itself quite selective. There are only occasional ventures into philosophical territory that Quine has not explored in his more academic writings, and these ventures seem to me rather disappointing. For example, in discussing the hoary topic of free will, Quine, like many other philosophers, favours the view that freedom is compatible with psychological determinism. Though all our actions and decisions may be completely caused, that does not necessarily deprive us of freedom. What matters is the way in which they are caused. One who adopts this view must explain what kind of causation allows for freedom, and it is in discharging this task that Quine seems to me to go astray.

David Hume offered the simplest account: we act freely when our actions are caused by our beliefs and desires. Yet it has seemed to numerous philosophers that this is far too simple, for there are individuals who are *compelled* by their desires — kleptomaniacs, alcoholics, compulsive smokers — and, though they satisfy their desires, they do not act freely. Another condition must be satisfied to ensure freedom, a condition that will differentiate those who are psychologically disabled from those whose *considered* desires are manifested in their behaviour.

Quine has no time for such complications. He offers the original Humean solution, and, in response to the possibility of "ill health of the offender's decision-making faculties", declares that psychological disability can provide no excusing grounds. This blunt treatment of the problems has the advantage of simplicity, but it surely errs in overlooking the existence of coercion from within as well as from without. Moreover, in explicitly frowning on the use of insanity pleas, it runs the risk of supporting much injustice.

Ultimately, though, questions of freedom, morality and punishment are not at the centre of Quine's thought. For many readers, *Quiddities* will serve as a superb introduction to central issues in contemporary thinking about logic, mathematics, language and science, an introduction that may lead them on to appreciation of Quine's half century of seminal writings. *Aficionados* will enjoy the witty reformulations of familiar themes and find a bonus in learning about the quintessential quirkiness of natural language. □

Philip Kitcher is a Professor in the Department of Philosophy, University of California, San Diego, La Jolla, California 92093, USA.

A particular way with words

Robert W. Cahn

The Art of Scientific Writing: From the Student Report to Professional Publications in Chemistry and Related Fields. By Hans F. Ebel, Claus Bliefert and William E. Russey. VCH: 1987. Pp. 493. Hbk DM98, \$59.95, £39.25; pbk DM48, \$24.95, £19.25.

PEOPLE who draft patents are apt to refer to the "state of the art", by which they mean current technique, technology, design, disposition, proportion, processing, assembly and the like. The book under review covers the state of the art of scientific communication, and going beyond that it also surveys the strategy and tactics of English style as practised by skilled (and unskilled) scientific communicators.

Books of this kind fall into two broad categories: reference manuals (such as the *Chicago Manual of Style* or *Copy-Editing — The Cambridge Handbook* by Judith Butcher) and discursive guides to good usage (such as *The Chemist's English* by Robert Schoenfeld). Although Ebel and his colleagues have contrived to steer a median course between these extremes, they do aim more towards systematic exposition and exhaustive instruction. They also seek to explain why as well as show how, and in these stated objectives they have succeeded.

The authors have a passion, not only for clarity and economy of style, but also for precision and consistency. Their attitude to both scientific writing and editing is reminiscent of the words of William Blake: "He who would do good to another must do it in Minute Particulars. General Good is the plea of the scoundrel, hypocrite and flatterer; for Art and Science cannot exist but in minutely organized Particulars". (Is that last full point in the correct place?)

The book treats, essentially, three broad themes: how to organize a report, dissertation, paper or book and how to achieve maximum clarity in the enterprise (the detailed advice on dissertations will be especially useful to anxious graduate students, who should find the paperback price accessible); the practicalities of drafting, correcting, typing or printing, drawing illustrations and chemical molecules, and proof-reading (the reasoned advocacy of computers as word-processing tools is particularly helpful); and the accepted conventions of citing the literature, using units correctly, setting out mathematical expressions, laying out tables, abbreviation of scientific terms, and chemical nomenclature. Although the

authors are chemists and pay special attention to chemists' difficulties, other readers merely need to do a little judicious skipping. In addition, there are 100 pages of appendices, mostly factual but also including a concise discussion of good English usage.

The authors certainly practise what they preach: I found only a single misprint and one grammatical solecism ("x is comprised of y").

On just one aspect of style the authors, uncharacteristically, blow both hot and cold: this is the vexed issue of the active versus the passive voice. While themselves preferring and occasionally using the active voice, they claim that their task is to describe the world as they find it and

that most of us prefer impersonal (and thus passive) constructions when describing our work. "The mixture was distilled" I accept. But I think that the phrase "It is thought that" is an abomination. We all of us, in writing about science, describe what was done and say what we think, and why. Putting opinions in the passive voice seeks to impose a semblance of objectivity on something that by its nature cannot be objective. As a former editor of the *Guardian* newspaper once (nearly) wrote: "Facts are sacred (and passive), opinions are free (and active)". □

Robert W. Cahn is in the Department of Materials Science and Metallurgy, University of Cambridge, Pembroke Street, Cambridge CB2 3QZ, UK.

Out of this world

Maxine Clarke

Sex and Scientific Inquiry. Edited by Sandra Harding and Jean F. O'Barr. University of Chicago Press: 1987. Pp.317. Hbk \$24.95, £19.95; pbk \$10.95, £8.75.

IT is difficult to guess at whom this book is aimed. From its title, one might imagine that it sets out to do the same job for women scientists that Germaine Greer did for artists. But although sex bias is one of the issues discussed here, the mish-mash of articles fails to do more than frustrate the reader.

All the contributions in this volume have been reprinted from *Signs: Journal of Women in Culture and Society*, and they address an enormous range of topics. The coverage is haphazard and arbitrary; thus, for example, Schiebinger's essay, in which the numbers of women scientists in various disciplines between 1928 and 1938 are analysed, stands alone, frozen in time, and is nowhere complemented by an account of how things have changed or how they might change in the future.

Lack of homogeneity is a common problem for editors of an anthology such as this. In tackling it, Sandra Harding and Jean O'Barr have restricted themselves to an introduction and, bemusingly, a footnote to two of the 15 articles in the book, setting out the authors' conclusions in a couple of sentences.

In the introduction, which by default is the only place where a clue to the philosophy of the collection might reside, Harding and O'Barr say the articles they have chosen reflect "five major focuses" of feminist concern in relation to science. These are: the social structure of science; uses and abuses of technology; bias; sexual meaning; and "epistemology and metatheory". Each of these topics is addressed by several authors, many of

whom use the style of writing in which several statements, each with a reference number, are strung together in a paragraph that generally ends with a take-home message along the lines of "chemistry sets help children develop different skills and aspirations than do Barbie dolls" (p.23). The clarity of argument in the essays is obscured almost totally by the use of jargon and of *ex cathedra* statements.

I would have found it easier to be sympathetic to the book if the authors had tried in some way to present alternatives rather than to analyse obscure, academic theses so exhaustively. Education, for example, is a subject barely touched on. An essay on the biological differences between men and women, and whether they are inherited or social, does not try to address what might be done to change attitudes. The ideological arguments of the essay are irrelevant in the light of the fact that many women have not begun to think about a scientific (or any other) career at the age of 15. But their interests would be well served by a more flexible education system that would allow them to take up science in their mid-20s, particularly those women who have suffered from class bias. Although the points discussed in the essay on sex differences may well be valid, they have less bearing on the problems that exist in the real world than other lines of argument might have done.

With a modestly priced paperback edition, the book is presumably intended for a general readership and, again presumably, aims to say something new to that readership. But with its disregard for the pragmatic and its espousal of the didactic, I imagine that it will be read only by feminists in academia and not by the browser in the bookshop. That is a pity, for what is needed is a book that is relevant to women scientists in the 1980s, not one that is stuck in the consciousness-raising era of ten years ago. □

Maxine Clarke is the News and Views editor of Nature.

Twist in the tale

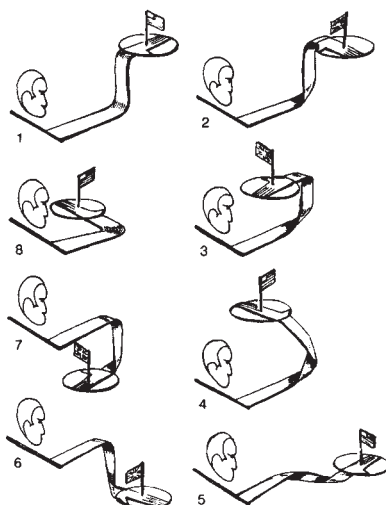
Clint McCrory

A Topological Picturebook. By George K. Francis. Springer-Verlag: 1987. Pp.194. DM78, \$33.

AFTER three decades of preoccupation with high dimensions, topologists have returned to the vivid world of surfaces and knots. An important factor contributing to the renaissance of visual topology has been the discovery by Bill Thurston of Princeton University of unexpectedly rich relations between non-Euclidean geometries and the topology of three dimensions.

But topologists are not very good at communicating their visions to others. Popular scientific writers still talk about turning coffee cups into doughnuts, leaving the fortunate impression that topologists merely belabour the visually obvious. What is needed is a modern version of the 1932 classic *Anschauliche Geometrie*, by Hilbert and Cohn-Vossen, with its emphasis on physical intuition rather than on abstract symbolism. George Francis's book is a step in the right direction. It is a manual on how to draw surfaces so that their topological complexity can be seen and understood.

Francis encourages the reader to learn to see by learning to draw. Many of his examples are entertaining visual brain teasers. Can you imagine a Moebius band whose border (edge) lies in a plane, or a two-sided surface whose border is a figure-eight knot? The book can be enjoyed simply as a picture book: each of its 87 illustrations tells a story using a number of related figures, and understanding these little picture stories is a pleasant and often challenging exercise. The last three



The plate trick — illustrating a geometrical principle popularized by Dirac, explained by Newman with the theory of braid groups.

chapters are much longer, and contain more intricate picture stories about important characters of modern topology: sphere eversion, mapping class groups and fibred knots (including some of Thurston's work). Through anecdotes, careful descriptions and references to

the literature, Francis skilfully unfolds the beauty of these subjects.

The foundation of the art of 'descriptive topology' is Hassler Whitney's theory of stable mappings of surfaces. The singularities (folds and cusps) of these mappings give visual clues to a surface's shape. (A corresponding theory of vision has been developed by the Dutch biological physicists Koenderink and van Doorn.) The topology of these singularities plays a fundamental role in René Thom's dynamical theory of catastrophes.

Computer graphics, the author contends, cannot yet compete with pen and paper in producing inexpensive, easily constructed images, and it is only in a postscript that Francis describes his initial forays into interactive animated computer graphics. I hope that he'll write a sequel on how to use a home computer to compose topological picture stories, for some of the most striking computer graphics of the past 15 years have been the result of mathematicians' attempts to see the objects they study. □

Clint McCrory is a Professor in the Department of Mathematics, University of Georgia, Athens, Georgia 30602, USA.

Disorderly conduct

Michael Berry

Chaos: Making a New Science. By James Gleick. Viking, New York: 1987. Pp.352. \$19.95.

IT is now more than a decade since it began to be widely appreciated that systems may behave randomly in spite of being governed by causal laws whose mathematical formulation is simple. Since then, the concepts and methods of chaosology have penetrated into virtually all branches of science. It is therefore surprising that we should have had to wait so long for a book-length elementary account of these developments, to complement the more philosophical and historical exposition by Prigogine and Stengers (*Order Out of Chaos*). And it is not a little shaming that the book should come from a journalist rather than one who is actively contributing to the subject.

The author has evidently taken a great deal of trouble to isolate the key ideas and explain them in simple terms. His technique is biographical: each concept is associated with one of the principal protagonists, whose character and personal history is interwoven with the science. The order thus imposed on the subject is not what would seem natural in a more didactic presentation — indeed the chapters could be arbitrarily permuted without losing clarity — but the immediacy and vividness provide some compensation.

Thus we are introduced to the mysteries of strange attractors through Edward Lorenz's early simulations of weather, Stephen Smale's mathematization of the stretching and folding of dough, David Ruelle's and Floris Takens's speculations on how fluids become turbulent, and the analogue computations of the 'Santa Cruz dynamical systems collective'.

Inevitably, the description of the central conception of fractals — geometric objects so wild that their dimensionality hovers between the integers — is enhanced by the larger-than-life personality of their creator, Benoit Mandelbrot. And the lonely, romantic genius of Mitchell Feigenbaum is pressed into service to emphasize his discovery that one of the ways in which chaos can develop is universal — that is, independent of the details of the system within wide limits.

Nobody should mistake this biographical approach for a serious history of the subject. It is a distortion to describe Laplace's determinism as "almost buffoon-like" just because that view has been superseded by the discovery of chaos. And the fact that Newton's theory of light eclipsed Goethe's theory of colour is no justification for calling the latter's obscurity "merciful".

Such inaccuracies are however unimportant by comparison with the unforgivable flaw in this book, namely its concentration on how the subject developed in North America, to the virtual exclusion of contributions by Soviet scientists. The seminal works of Kolmogorov



The Whitney bottle — recipe, middle, for the 3-dimensional shadow, right, of a Klein bottle in 4-space from 3-dimensional sections.

NEW BOOKS

Mammals

Their Reproductive Biology and Population Ecology

J.R. Flowerdew

Mammals display a wide variation in their reproductive biology and life histories as well as in size, shape and behaviour though united by many common characteristics. The differences observed between the various species indicate adaptation to particular evolutionary pressures.

This book offers the facts and considers the theories in the light of experimental evidence, illustrating, with examples of physiological and behavioural adaptations, the links between mammalian reproduction and population ecology.

£13.50 net paper

240 pages 107 line and 3 half-tones

Publication October 1987

ISBN 0 7131 2896 8

Plankton and Fisheries

John Grahame

The organisms of the plankton are responsible for the production and transfer of most of the food energy available in the sea. Effective utilisation of fish as a food resource for man demands an appreciation of plankton ecology and a variety of biological factors – both physiological and ecological.

The author has drawn together the pertinent aspects of these diverse subjects, relating the fields of oceanography, plankton biology and fisheries biology, and making useful reference to ecological literature outside marine biology.

£7.95 net paper 144 pages

67 line drawings and 3 half tones

Publication October 1987

ISBN 0 7131 2941 7

Subtidal Ecology

New Studies in Biology

Elizabeth Wood

This book provides a 'modern' approach to sublittoral marine ecology, an area of great interest to pollution and oil field developments. The author covers geomorphology and hydrography, the distribution of the flora and fauna and their adaptations to various substrata, the interrelationships of the communities, food webs, and discusses the qualitative and quantitative techniques and methods of studies.

£4.95 net paper

224 pages illustrated

Publication October 1987

ISBN 0 7131 2957 3

Edward Arnold

a division of

Hodder and Stoughton

41, Bedford Square

London WC1B 3DQ

and Arnold (and also of Moser in the West) on the stability of the Solar System, of Sinai, Alexeev and Pesin on fully chaotic motion, and of Chirikov on criteria for chaos in accelerators and plasmas, barely rate a mention. This amounts to a reconstruction of history worthy of the Kremlin in the days before *glasnost*. It is hard to fathom the reason, especially as a reading between the lines makes it appear that many of those whom the author interviewed must have fully and generously acknowledged their indebtedness to colleagues from the East. I can only guess that the width of the Atlantic, the height of the Iron Curtain and the terseness of some Soviet scientific writing combined to hide the fact that Russians have personalities too – and in this style of writing a lack of 'human interest' seems to mean no interest at all.

Apart from this unpleasant chauvinism there are a few minor errors. Friction is confused with nonlinearity, and there is the curious assertion that "Physicists had theorems, mathematicians had conjectures". In the main, however, the explanations are clear and accurate, and the intensely visual appeal of the subject is reflected in diagrams and colour plates. There is a useful bibliography to guide those who wish to dig deeper.

On balance I recommend this book as a lively introduction, conveying to non-specialists the intellectual excitement accompanying these new insights into the relation between the simplicity of the laws of nature and the complexity of the world as it is. □

Michael Berry is a Professor in the H.H. Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, UK.

A massive problem

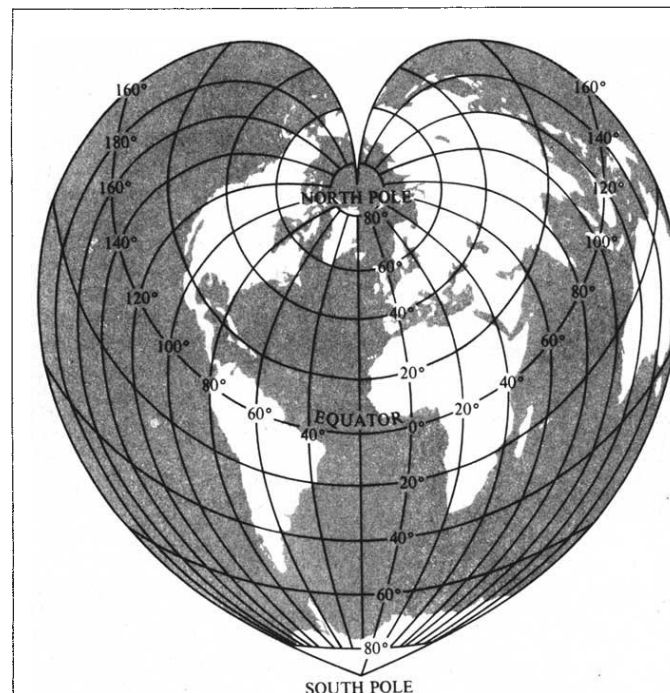
Robert H. Sanders

The Omega Point: The Search for the Missing Mass and the Ultimate Fate of the Universe. By John Gribbin. *Heinemann: 1987. Pp. 245. £12.95. To be published in the US in January by Bantam, pbk \$9.95.*

THERE are a number of compelling reasons for believing that the Universe can be described by the flat space Friedman solution to Einstein's field equations. This requires that the average density of matter in the Universe should be very nearly equal to a critical value; in other words, that the density in terms of this critical density, designated "omega", should be unity to high precision.

In this highly readable book, John Gribbin reviews theoretical and philosophical arguments in favour of $\Omega=1$ for the non-specialist, and the efforts by astronomers, cosmologists and particle physicists to solve the omega problem – the question of whether the Universe is open and will expand forever or is closed and will eventually recollapse.

Gribbin argues convincingly that the only natural value of Ω in the standard cosmological models is unity; the work on quantum gravity by Stephen Hawking provides a theoretical basis for expecting Ω to be very nearly unity as an initial condition, and modern theories of the unification of forces provide a mechanism, "inflation", for smoothing and flattening the Universe. He goes on to stress that if we accept that Ω is unity, then astronomical observations are telling us that



Map-makers must resort to trickery to depict a round Earth on a flat sheet of paper. Johann Werner's heart-shaped map, popular in the 16th century, represents most of the main land masses accurately with little distortion in shape.

The illustration is reproduced from the chapter "Curious Maps" in the inimitable Martin Gardner's latest collection of puzzles. The book is entitled *Time Travel and Other Mathematical Bewilderments*, and is published by W.H. Freeman. Prices are hbk \$17.95, £16.50; pbk \$12.95, £11.50.

most of the mass of the Universe should be non-luminous matter. The visible matter in galaxies accounts for less than 1 per cent of the critical density. The greatest predictive success of the standard Big Bang cosmology, the nucleosynthesis of the light elements, implies the density in ordinary baryonic matter could be at most about 10 per cent of the critical density. Thus the dark matter must be non-baryonic and may consist of one of the undiscovered particles predicted by unification theories such as supersymmetry.

Up to this point the discussion is logical and persuasive. Gribbin succeeds in convincing the reader that, in the context of modern cosmological theory, the *a priori* arguments for $\Omega=1$ are overwhelming. However, when it comes to the empirical evidence for $\Omega=1$ he gives the impression that astronomers have all but wrapped up the case in favour of a flat space Universe. The fact is that, in spite of all the hopes for $\Omega=1$, the astronomical evidence is far from clear cut. In gravitationally bound systems — from single galaxies to rich clusters of galaxies — there is no indication that the dynamically measured mass is much more than ten times the visible mass. This is completely consistent with all dark matter being in baryons and with $\Omega \approx 0.1$. It is possible, as the author notes, that the dark matter is spread much more evenly throughout space than the visible matter (a result of "biased" galaxy formation), but the recently discovered large-scale streaming motions in the Universe are not consistent with this idea.

In his enthusiasm Gribbin goes further and practically certifies WIMPS — weakly interacting massive particles — as the dark matter closing the Universe. The principal evidence given for this is the suggestion that the presence of such particles in the Sun can account for the anomalously low detection rate of solar neutrinos. Although this is an intriguing possibility, there remain outstanding theoretical and observational objections. This is only one example of the essential shortcoming of the book: ideas or suggestions that are only tentative are presented as firmly established and all but proven.

In a popular account of a rapidly developing field, an optimistic and enthusiastic style is justified by the need to convey the excitement of research at the frontier. But the sense of excitement should not be achieved at the expense of accuracy. The work described here is a great story of human enquiry into some of the most fundamental questions in science. It is intrinsically exciting, not only because of the puzzles that have been solved, but also because of those that remain. It is a mistake to exaggerate the successes and downgrade the mysteries. □

Robert H. Sanders is at the Kapteyn Astronomical Institute, Postbus 800, 9700 AV Groningen, The Netherlands.

Seeking asylum

Michael Shepherd

A Social History of Madness: Stories of the Insane. By Roy Porter. Weidenfeld & Nicolson: 1987. Pp. 261. £14.95.

A PROMINENT medical historian has pointed out that

Until relatively recently, the history of psychiatry continued to be written primarily by practising psychiatrists, and reflected to a large extent the professional preoccupations of its authors. . . . Increasingly, however, the history of psychiatry has not been the exclusive domain of the psychiatrist, as social, intellectual and legal historians, sociologists, criminologists, anthropologists, philosophers and historians of medicine, among others, have examined many cultural and historical aspects of mental disorders [W.F. Bynum, in his contribution to *Handbook of Psychiatry*, Vol. 1; Cambridge University Press, 1983].

As might be expected, many of these newcomers bring a very different perspective to bear on the field, often deriving

purpose. The bulk of the volume, accordingly, consists of the presentation of a series of biographical sketches of mad people, some world-famous and others obscure. These are related, chapter by chapter, to several broad topics — power (George III, Christian VII, James Tilley Matthews), artistic genius (Robert Schumann, Vaslav Nijinsky, John Clare), religion (William Cowper, George Trosse, Christoph Hartzmann) and so on. Much of the material is familiar but the text reads trippingly and an appendix of annotated suggestions for further reading indicates the extensive background from which the information has been drawn.

The author's declared intention — "simply and quite literally to see what they had to say" — is, however, economical with the truth. In his opening chapter, "Madness and Psychiatry Talking" he reveals the underlying relativism of his approach, contending that the madman's eye-view of his own language, metaphor and culture makes its own sense, and that "madness may have moved towards psychiatry . . . we may be witnessing yet



Mary Evans

Inner visions — power, art and religion represented by George III, Nijinsky and Cowper.

their inspiration from Michel Foucault's seminal volume, *Madness and Civilisation*, which goes some way towards challenging the legitimacy of the whole medico-psychiatric enterprise. As one of them, C. Unsworth, has remarked:

Psychiatric history looks different if psychiatric medicine's status as a science is not assumed, but treated as a product of historical development, with focus upon the historical contingency of psychiatric ascendancy, the structure of psychiatric discourse, the strategies of the medical profession in laying claim to the psychiatric territory, and the specific contribution of the psychiatric apparatuses and practices to the perpetuation and development of socio-political orders [*The Politics of Mental Health Legislation*; Clarendon, 1987].

Roy Porter, a practised and knowledgeable social historian, exemplifies the outlook in this book. He presents the thoughts and feelings of some two dozen individuals regarded as insane at various times, drawing principally on their autobiographical writings supplemented by literary and historical sources for the

another mode of folie à deux, madness and psychiatry as doubles".

This theme is most clearly illustrated in the penultimate chapter, entitled "The Therapeutic God", which refers to four individuals. Three of these are Sylvia Plath, John Butt and one of Sigmund Freud's best known cases, the so-called 'Wolf-Man'. The fourth is Freud himself, whom Dr Porter puts alongside his other subjects as "one who underwent a mental disturbance and who subsequently told his life story". The boundaries between doctor and patient, between reason and unreason, are blurred by the portrait of a straw man: the founder of psychoanalysis represents the thaumaturgy not the science of the psychiatric tradition. Medical history, as Henry Sigerist always insisted, is not only history but medicine as well. And psychiatry is part of medicine. □

Michael Shepherd is Professor of Epidemiological Psychiatry at the Institute of Psychiatry, University of London, Denmark Hill, London SE5 8AF, UK.

All in the mind?

John C. Marshall

The Oxford Companion to the Mind. Edited by Richard L. Gregory. Oxford University Press: 1987. Pp. 856. Hbk £25, \$49.95.

FROM "abacus" to "Zeno of Elea", by way of "masochism" and "metempsychosis", this massive compilation really does contain something for everyone. In emulation of that well-known psychologist Scheherazade, *The Oxford Companion to the Mind* contains 1,001 entries. These range from brief one-liners ("Analogue. See Digital") to substantial essays on asylums, colour vision and table manners. The contributors (and the distinguished folk they discuss) are a similarly motley crew. Psychologists, physiologists, philosophers, psychiatrists and pedagogues take the lions' share, although anthropologists and anatomists, computer scientists and criminologists, linguists and logicians, neurologists and neurochemists all contribute to the feast... to say nothing of the television astronomer and one or two theologians.

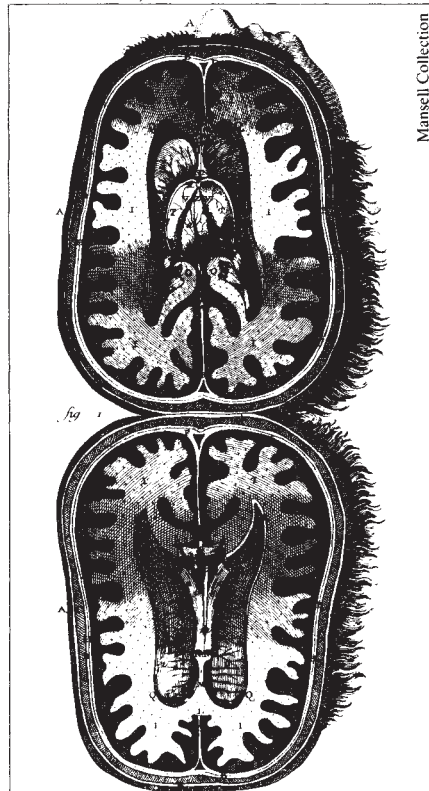
Clearly, then, Richard Gregory's *Mind* is a catholic Church, and one dreads to think of the amount of time and patience that he has devoted to the project. Indeed the decade or more required to organize the 200 or so contributors is such that quite a number of them have died in the interval between writing and publication. Nevertheless, a reviewer must ask whether the effort was worth it. There is, of course, no point in nit-picking; with an enterprise of this scope there is no way in which an editor can please all of the people all of the time. But one can ask to whom this Oxford *opus* will be a companion, and one should worry about the overall picture of mental life that is being presented.

Unsurprisingly, the best articles are those that discuss a substantive body of sober scientific work. Many of these are concerned with the physical substrate of mind in the central nervous system. Thus Colwyn Trevarthen (on brain development), Marcus Raichle (images of the brain in action), Alan Cowey (neuro-anatomical techniques) and H.F. Bradford (neurotransmitters and neuromodulators), for example, all contribute useful summaries for the serious student while still allowing that strange creature the 'general' reader to follow what's going on. And there are many other more 'cognitive' articles that are equally good.

Such 'core' topics as language, perception, memory and problem-solving are also well-covered; to pick just a few of the many fine essays, Noam Chomsky writes lucidly on the growth of language within the individual mind/brain; Bela Julesz

outlines his theory of (pre-attentive) visual processing as an early warning system; J.A. Deutsch summarizes experimental approaches to the study of learning and memory; and Peter Wason explicates various seemingly trivial logic problems that fox bright graduate students (for non-trivial reasons).

The numerous thumbnail sketches of the great and the good are similarly helpful for anyone who has forgotten (or never knew) that Charles Babbage, who built the first programmable computing machine, also invented the ophthalmoscope; that Golgi (Camillo) was a Lombardian, as well as a stain for



Open-minded — an 18th C view of the brain.

neurons; or that Arthur Schopenhauer was "a lonely and unloved bachelor, befriended only by his poodle Atma". Many of the more informative biographical entries (including those on Kurt Goldstein, Heinrich Kluver and Wilder Penfield) were contributed by Oliver Zangwill, the doyen of British psychology, whose recent death was the loss of a scholar and a gentleman.

The *Companion* contains no entry between "Platonic forms" and "pleasure centres", the point where one might have expected to see reference to that most serious of all human activities, play. Nonetheless, aesthetic pleasure does find a place in *The Mind*. The late Richard Jung, a neurologist from Freiburg, has contributed a beautiful and insightful little chapter on art and visual abstraction in which he relates graphic marks to the types of visual representation computed by different physiological mechanisms in the

brain; Natasha Spender, a concert pianist, provides a similarly exciting piece on the psychology of music. This article expertly covers music as psychoacoustic stimulus, as perceptual pattern, as formal language, as sensory-motor skill and as emotional experience. A 'reductionist' Natasha Spender is not.

The *Companion* does, however, contain an essay on reductionism in psychology, albeit not by a philosopher. Rather, the topic is discussed by the (late) Soviet neuropsychologist Alexander Romanovich Luria, who approaches it as a good Marxist: "The explanation of the phenomenon is supposed to lie not in its reduction to single elements but rather in its inclusion in a rich net of essential relations". As Karl said: science should ascend to the concrete, not the abstract. I was grateful for Luria's offering because the philosophers themselves make heavy weather of their supposed contributions to clarity and knowledge in this volume. True, the entries on the classics (Aristotle, Descartes, Plato, Spinoza, for example) are clear and often entertaining; but for the rest there is far too much second-hand 'analytic' philosophy of the sort that clutters up the shelves in academic bookshops. Philosophers of physics, say, are expected to know some physics; why is it that philosophers of mind seem to feel that an afternoon of quiet meditation is a sensible alternative to learning what has actually been discovered about the structure and functions of the mind/brain?

In his autobiography, G.E. Moore wrote: "I do not think that the world or the sciences would ever have suggested to me any philosophical problems". Likewise, Ludwig Wittgenstein (in his notes on *Culture and Value*): "At bottom I am indifferent to the solution of scientific problems". One suspects that the twentieth-century 'professionalization' of philosophy is, in part, responsible for this sorry state of affairs. Be that as it may, Richard Gregory could have found more philosophers who were prepared to do their homework.

The more theological entries are somewhat better, although a little skimpy given the important role that religion has always played (for good and ill) in humanity's conception of itself. The Reverend O.J.W. Hunkin's piece on religion is calm and thoughtful, and is neatly complemented by Sir Edmund Leach on humanism. Eastern approaches to the mind are particularly well represented, and variations upon Islam abound with copious entries for teachers whose names begin with Al- or Ibn. Interesting as these pieces are they do not help us to understand the state of mind that currently reigns in the Gulf (and may yet send us all to perdition). Christians and Jews have a leaner time in the *Companion* — no Hildegard of Bingen or Nachman of Bratslav graces these pages.

Indeed the only Western mystic we find is that Viennese engineer who was "indifferent to the solution of scientific problems". I hope that the propeller Richard Gregory reports him as designing actually worked.

But the very worst I have left 'til last: the *Companion* has entries (some of them quite long) under astrology, extra-sensory perception, ghost, levitation, paranormal (this, that and the other), spiritualism and telekinesis. I do realize that many people are keen on such things, but I see no good reason why that attitude should be

pampered. Just as the first responsibility of a hospital is to not spread disease, so the first responsibility of a *Companion to the Mind* is to not encourage superstition. There are plenty of problems awaiting solution before we need concern ourselves overmuch with mysteries. I would happily have swapped all these para-chapters for an extended essay by that truly great student of matter-over-mind, James ("The Amazing") Randi. □

John C. Marshall is in the Neuropsychology Unit, part of the Neuroscience Group at the Radcliffe Infirmary, Oxford OX2 6HE, UK.

On speaking terms

J.Z. Young

Mindwaves: Thoughts on Intelligence, Identity and Consciousness. Edited by Colin Blakemore and Susan Greenfield. Basil Blackwell: 1987. Pp. 525. £19.50, \$24.95.

RECENT discoveries in neuroscience have introduced new controversies into the problem of how best to speak of the relations between mind and brain. This book provides 32 opinions on the subject and makes a fascinating study of the different approaches of philosophers and scientists. Neuroscientists feel that they have been helped by adopting part of the terminology of communication engineering and computer science, but some philosophers claim that the use of such terms introduces dangerous conceptual confusions into the debate.

The philosopher Peter Hacker and linguist Roy Harris have taken their task to be correction of the mythological language used by physiologists, picking on Frisby, Barlow, Phillips, Zeki and especially as expressed in my own book *Programs of the Brain*. They object to such terms as 'information', 'coding', 'language', 'mapping' or 'representation' being used in genetics and neuroscience, saying that they are misapplied and may lead to "disastrous equivocations" or misleading metaphors that cannot be tested.

All discussion is helpful but these criticisms show surprisingly little sense of history and are short-sighted and obscurantist. They miss the point that all living things are agents and communication systems. Biology advances in line with technology, applying to human beings and animals the understanding and terminology derived from prosthetic devices that replace or augment our functions. These terms from communication engineering are part of current language. Geneticists, biochemists and neuroscientists find it useful to speak of communication of information by codes and the setting up of representations: most of the contributors

to this volume use such words and they are not likely to stop doing so because philosophers insist that these terms should be used only in restricted senses and about whole human persons.

It is worthwhile examining this problem further. The ability to communicate by conventional code signs is characteristic of living things. It provides the information that maintains the pattern of order that allows life to delay the increase of entropy. Signals that transmit this semantic information are indeed parts of the pre-arranged communication systems of genes, hormones and brains. To quote Hacker himself: "A symbolic description is presumably an array of symbols which are so combined as to yield a true (or false) characterisation of a certain aspect of the world. It must be cast in a certain language which has a vocabulary and a grammar". Living things can be said to use symbols in just this way and the points of view of these philosophers obscure this fundamental fact about life. Moreover, the critics imply an unacceptable dualism: they seem not to realize the axiom of psycho-physical intimacy; you and your brain are inseparable, why separate them linguistically? Hacker says that "it makes no sense to speak of the brain containing knowledge or information written in its own language". On the contrary, neuroscientists are beginning to give good sense to that statement. Jeffrey Gray has no doubts: "all human languages are stuffed full of rules. . . . The rules, then, must be contained in the heads of those speakers and hearers".

The linguist Harris might ask "Which rules?", because grammars are embarrassingly numerous. How do you hope to find them in the brain when linguists do not know what they are? One must agree that scientists when speaking of a 'language' do not consider all the subtle problems that have arisen since de Saussure discussed the difference between 'langage' and 'langue'. There are difficulties in the use of such terms as 'information', 'coding' and 'representation' and we try to define them in general biological terms, as I have done more recently in my book *Philosophy and the Brain*. It is unsurprising

that such changes of use upset those who do not look at the whole living world, but it is disappointing that the critics show so little understanding when one hoped for help with common problems.

Surely a wider view will help with the classic problem of the mental and the physical. Harris believes that neuroscientists "do not wish to be restricted by accepting the traditional distinction between 'brain' and 'mind'" and he claims that the motivation of the biologist is that he "wants to be in on the act when it comes to solving these ultimate scientific mysteries: how do human beings think? and what is thought?". This is an unworthy tone for one scholar to adopt of others who are working in an obviously related subject. Most neuroscientists do indeed feel that a re-consideration is necessary. In a wise article, Paul Seabright discusses the "tensions that arise between the mind we experience and as explained by science". For one thing "the physical world is everywhere; it has no pockets into which non-physical objects could fit. . . . But most of us also think many things that exist are not *just* physical objects". So how are we to describe this property of mentality that attaches to ourselves and perhaps at least to some animals?

The first essential is not to separate the aspects too widely. Stephen Clark echoes Seabright on this: "a surprising number of people still seem to think that offering an explanation in terms of intention and feelings or the like is somehow incompatible with an explanation in terms of physiological conditions, as if intention occupied space to the exclusion of nerve fibres". Ted Honderich presses the point that "Neuroscience . . . must change the philosophy of mind . . . mainly because it establishes the axiom or proposition of psycho-physical intimacy". He goes on with his suggestion of how we should formalize this co-occurrence: "Consciousness, truly described by the metaphor of the interdependent subject and object, does exist". He characterizes their relationship as one of nomic (law-like) connection. Nicholas Humphrey uses a different metaphor: "consciousness is in fact a picture of the workings of the brain".

Euan Macphail, a psychologist, asks "When we solve problem X, what processes are taking place in our minds (or brains)?" This has the unfortunate effect of suggesting that the mind is somehow an entity that can contain processes. Surely, as Ayer has said, "We do not need to conceive of minds as substances, or indeed as entities of any kind". It is better to say with Colin McGinn, "Still, the brain is a physical entity and it is conscious, so it must have *some* design feature, presumably 'physical' in nature (whatever that might mean), that *makes* it conscious". But, he continues, "we do not at present know what that feature is, and so we do

not at present know how to build a machine which matches this achievement of the brains".

John Searle's explanation of the 'causal' relationship between mental states and the brain is "The existence of two causally real levels of description in the brain, one a macro-level of mental neurophysiological processes and the other a micro-level of neuronal physiological processes . . . exactly analogous to the existence of two causally real levels of description of a hammer". One wonders whether the words "exactly analogous" suffice to solve this ancient question?

Jack Eccles provides his frankly dualistic solution with the concept of mental-neural event neurons (MNE) sitting at higher levels above mere neural event neurons (NE): "In special circumstances their [MNE] firing would be in unison (identity) with mental events". Searching

for the means of such mental-neural interaction, Eccles calls upon the possibility suggested by Margenau that the mind operates as a "probability field of quantum mechanics". However Seabright points out that quantum laws are random and so cannot be the basis of psychological law or intention, a view that is held also by Honderich. Roger Penrose too considers that "the introduction of quantum mechanics [makes] no essential difference", but he raises the interesting suggestion that "there is something very incomplete about quantum theory as it stands".

Rodolfo Llinás makes a determined effort to show how "sensory inputs [are placed] 'into internal context' . . . so that these inputs may be *understood* by the nervous system". His account involves "organised patterns of electrical activity in neurons, which may be described and formally treated as implementing a

functional space". He ends by hoping that "Ultimately understanding of the mind may not be as intricate as our vanity hoped or our intellect feared".

Other authors discuss the possible functions of mentality and consciousness, and several of them are ready to ascribe these properties to animals. Nick Mackintosh provides evidence that even "rats and pigeons in boring experiments on conditioning" show evidence of knowledge of and belief about the world, but he warns that "this is not necessarily an appropriate language for scientific analysis". Herbert Terrace argues that in apes we have evidence of animal thought but not the characteristics of human language. Further, "recent experiments on animal memory provide compelling support for the idea that animals do form representations". Evidence for internal representation in fishes is also provided by Macphail from studies of their electrical emissions.

Thus the contributors to this book do not seem to wish to dismiss mental events; but they do try to clarify their relations with brain events. Many of them write of the brain as using a language of symbols and setting up maps and representations. It is difficult to see that this is leading them to the confusion over the working or nature of the brain which Hacker and Harris fear. Of course, these words are not used in exactly the same sense when applied to brains and human language. But because all mentality is closely correlated with brain activity, it is unlikely that wholly different terminology is needed for the two. Perret, Rolls and Caan have shown that there are cells in the cortex of a monkey that are selectively responsive when the face of a monkey or human is shown. Surely the presence of these cells in monkeys, and no doubt in man, contradicts Hacker's claim that "nothing in the cortex constitutes a 'symbolic representation' of the creature's environment".

We need techniques to study how such gnostic cells interact with others to provide the action system that allows a monkey or a man to have mental and physical reactions appropriate to friend or foe. Such an enquiry involves considering the whole brain as an agent, which is a holistic approach that is not easy for physiologists. Objections to the use of the metaphors under discussion can actually delay discovery by forcing investigators back to the use of simple physical principles only. So I hope that neuroscientists will not be put off by philosophical sophistries. We need to pay more attention to human and linguistic models, not less, in order to help with the task of understanding mentality and the brain. □



Mental efficiency — a soldier's view of the US Army's psychological examinations, first featured in the Camp Sherman News of 1918. When Stephen Jay Gould gave similar tests to Harvard undergraduates, many did well but some barely reached 'C' level. The cartoon is taken from *A History of Psychology* by Thomas Hardy Leahey. Publisher is Prentice-Hall, price is £36.40.

J.Z. Young, Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD, UK, is Emeritus Professor of Anatomy in University College London.